

# nature

## HOW TO SURVIVE THE RECESSION

Science and the  
global squeeze



**FROM GENOMICS**  
A questionable heritage

**MRI SCANNERS**  
How were fMRI detectors

**EVOLUTION**  
Complex joints unevolved

nature  
ISSN 0028-134X

## Collective responsibilities

China should stop discouraging scientists from setting up learned societies.

The spread in China of unproven stem-cell therapies for conditions such as epilepsy and spinal cord injuries has left the nation's health authorities concerned. There is no clear evidence that these treatments work — nor that they are killing people. Of the thousands of patients from China and abroad who have been treated, some seem to think that they have been helped, even if only modestly, and many more are ready to fork out thousands or tens of thousands of dollars to try out the treatments. Are the clinicians taking advantage of people desperate for a cure? How can the government — and the potential patients — make sense of this?

One obvious place from which to seek guidance would be the national stem-cell society. But China doesn't have one. A group of scientists, including many of the country's most prominent and internationally established researchers, are trying to create one.

The Chinese authorities, however, tend to have an aversion to congregations — especially those such as the Falun Gong, which they believe pose a threat to the country's stability. So the Ministry of Civil Affairs keeps a tight hold on who is allowed to organize in any formal sense. As a result, China's stem-cell hopefuls must go through the slow process of planning and applying to become a 'level 2' society. That means they have to convince an established society to take them on as an appendage, which will dramatically reduce their ability to function effectively. A level 2 society doesn't control its own purse strings and decisions have to pass through the parent organization.

Yet, as the example above illustrates, allowing scientists to draw together can only benefit China, both by helping scientific progress and by assisting with the challenges faced by the Chinese nation.

It is not just the government that needs to rethink its approach: the researchers themselves need to pursue newer forms of social

organization. Scientists in the south often don't know what is happening in the north and vice versa. Most of the current learned societies do not function well. Annual meetings are often a matter of pomp, with elite researchers showing up to swagger about and form cliques based on pedigree rather than scientific views. Introducing graduate students to the broader community is a low priority. Constructive criticism is more likely to be taken as grounds for breaking off relations than as insightful advice. Many scientists simply don't bother to go.

Sometimes 'megaprojects' draw researchers together. But the planning meetings for such packages can be more like dividing the spoils than building the most constructive research programme.

China's science loses competitiveness because of these failings. Stronger societies would pave the way for better communication and more productive collaborations, and would allow a platform for feedback of scientific criticism. That, in turn, would provide a body of honest reviewers with whom funding bodies could consult. Too often, instead of listening to a variety of voices to get a representative view from 'the community', funding bodies listen only to certain well connected scientists. Strong domestic scientific societies have the additional benefit of being reference points for constructive contact with scientists and societies elsewhere. And they can also act as advisory bodies to the government.

Gone are the days of small research communities in China. Science has grown significantly, to China's credit and benefit. For the country to benefit more fully, networking by its researchers likewise needs to be allowed to flourish.

**"Scientists in the south of China often don't know what is happening in the north and vice versa."**

## Identity crisis

It is time for all involved to tackle the chronic scandal of cell-line contamination. Funders first.

Some 40 years after it was first recognized, the use of contaminated and misidentified cell lines in biological research remains a growing problem. But it is a problem that has a simple solution: routine, cheap, DNA profiling of laboratory cultures. It is now time to implement that solution. To do so, scientists need the funding and motivation to verify the cell lines in their possession, as well as a curated electronic database of authenticated DNA profiles against which they can compare their results.

Thousands of biology labs use cell lines, yet many do not know that between a fifth and a third of the lines in common use may not be what they seem. In the past 25 years, numerous studies, as well as the

experience of cell-culture repositories in the United States, Britain, Germany and Japan, have found that 18–36% of cultures contain a misidentified species or cell type. The effect of using such cells varies depending on the project involved. When the lines are used as a source of biochemicals, for example, the misidentified lines are innocuous. Deployed in the study of a general cellular process, they can have minor drawbacks. But on the rare occasions that the cell lines are thought to reflect the properties of a particular tissue, cancer or disease state, the outcome can be severely damaging as funding and research get driven into work based on false premises.

To make matters worse, papers are still published that contain unwarranted conclusions derived from misidentified lines. It is ironic that many researchers who are obsessed with using only the highest-quality chemicals and biologics from the most trusted suppliers don't think twice about using cell lines known to be misidentified.

Cell repositories do carry out quality-control assays on deposited lines, although the tests performed vary. Even a venerated panel of

60 tumour lines at the US National Cancer Institute was found to have some that were either HeLa (the first human cancer cell line) or subcultures of one another.

This crisis of identification can be solved by analysing repository cell lines using DNA fingerprinting — short tandem repeat (STR) assays — and making the ‘authenticated’ profiles available in a database. Some of the cell-line profiles in the American Type Culture Collection, for example, already have their STR profiles listed. The German DSMZ cell repository performs DNA profiles for every line, but has also reported that 29% of its human tumour line deposits are cross-contaminated. It costs between \$20 and \$400 to fingerprint a cell sample (depending on country and circumstance), and some predict that the \$2 STR analysis is not far away. At that price, what lab could not afford to regularly recheck its cultures?

In an open letter in 2007, Roland Nardone of the Discovery Center for Cell and Molecular Biology in Washington DC, and his colleagues brought the issue of misidentification to the attention of Michael Leavitt, until recently director of the US Department of Health and Human Services. This moved the issue forward: by the end of the year the National Institutes of Health (NIH) formally recognized in a public notice that “misidentification of cell cultures is a serious problem”. However, the notice went on to state that “it would be impractical

for the NIH to require application of particular methods in all grant applications”, and put the onus on peer reviewers to quality-control their colleagues’ research proposals and manuscripts.

This merely capitulates to the status quo. Four decades after the problem came to light, it is time for this cavalier attitude to be jettisoned. Repositories need to authenticate all of their lines, and the NIH and other major funders must direct support accordingly. The STR profiles should be lodged in a global database that provides tools for readily comparing a culture’s fingerprint with authenticated profiles. The funders should motivate investigators by encouraging the inclusion in grant proposals of expense estimates for cell-line verification, in recognition that this quality assurance will increase the costs of research. The community, in turn, should accept that it makes sense to verify cell lines routinely.

Once this research framework is sufficiently established, major funders will be able to require the validation of all immortal cell lines in order for investigators to retain funding, and journals should (and *Nature* will) require that all lines used in a paper were verified before publication. ■

**“Repositories need to authenticate all of their cell lines, and the major funders must direct support accordingly.”**

## Overhead hazards

How to keep Earth orbits usable.

“*S*pace is big,” the British humourist Douglas Adams once observed litotically. “I mean, you may think it’s a long way down the road to the chemist, but that’s just peanuts to space.” And yet, in all this vastness, there are now some regions so crowded that it is possible for a pair of satellites no bigger than compact cars to collide by purest accident, with no malice aforesight, as happened 800 kilometres above Siberia on 10 February (see page 940).

The fact that humans have managed to spread traffic hazards beyond the confines of their planet is humbling. It is also a serious problem — one that could severely hamper scientific research, weather forecasting, commerce and the national defence of various nations. It needs to be sorted out.

There are two complementary ways of doing this. One is for all satellite operators to abide by debris-minimizing rules such as those promulgated by the UN Committee on the Peaceful Uses of Outer Space: depressurize fuel tanks when you are finished with them, take steps to make sure that batteries don’t explode, shut down flywheels after the mission is over and, most crucially, drive low Earth orbit satellites to a fiery atmospheric death when they have fulfilled their tasks. The fact that there is no clear way to enforce such practice does not mean that the international community should not try to insist upon it.

The other response is better tracking, which would allow satellites at risk to manoeuvre out of the way of each other. The military establishments in both the United States and Russia track objects in

orbit. The Americans make some of their data available to the world at large; the Russians, to their shame, do no such thing. But the shared US data, although better than nothing, are more crude than those that the military keeps to itself. Better data would allow better decision-making by satellite operators weighing evasive action.

The US military has various reasons for not providing the very best data to all who ask for them. To do so would reveal the capabilities of the US surveillance systems — and perhaps their blind spots — in uncomfortable detail. It would also make the targeting of anti-satellite weapons easier.

One solution would be to release the data to a trusted intermediary with the analytical power to look for potential collisions and alert operators when things look bad. Another would be for the US military to do something along these lines itself. It already provides such services for some high-value missions by NASA and some allies. Expanding the service would make sense. If the US national security apparatus were to reduce the number of future collisions by letting third parties know of the risks, it would be improving the survival chances of its own spacecraft as well as everyone else’s — and no one has more valuable assets in low Earth orbit than America’s soldiers and spooks.

In the long run, an independent tracking system with its own data sources would be the ideal solution, and to its credit Europe has made some moves towards developing such a thing. But on the world stage this does not seem to be a priority. The problem is that by the time it becomes one — maybe two or three collisions down the line — the threat may have been ratcheted up far enough to be considerably less tractable. Every time two objects in orbit collide, they create more debris that can lead to more collisions. The way things are going, this will be one of those problems where the need for action becomes truly obvious only after it is too late. ■

# RESEARCH HIGHLIGHTS

## On the cusp

*Geology* 37, 187–190 (2009)

They might look like the footprints of giants or of alien visitors, but the quasi-geometric shapes adopted by some elongated lakes or ponds have a natural explanation.

The shores of such water bodies are washed by high-angle waves, which make them prone to instability caused by erosion and deposition of shoreline sand or gravel, say Andrew Ashton of the Woods Hole Oceanographic Institution in Massachusetts and his colleagues. Their simulations show that this process creates cusp-shaped capes and spits, and that the effect of these on wave patterns often leads to cusps on opposite shores ‘attracting’ one another. The cusps ultimately link up, forming bridges that divide the water body into a series of smaller, often oval, lakes.



NASA/DIGITALGLOBE/TERRAMETRICS/GOOGLE EARTH/A. ASHTON

## ASSISTED COLONIZATION

## Flutter further north

*Conserv. Lett.* 2, 45–51 (2009)

Many species are edging northwards in response to climate change, among them the UK-resident butterflies marbled white (*Melanargia galathea*) and small skipper (*Thymelicus sylvestris*). In 1999 and 2000, Stephen Willis of Durham University, UK, and his colleagues collected several hundred adults of each species and moved them many kilometres north of their current ranges.

Both butterflies thrived in their new homes, suggesting that their ability to disperse northwards is slower than the rate at which suitable habitat is becoming available owing to climate change. The authors say that actively moving certain species may help them to survive climate change.

## ECOLOGY

## Perils of monoculture

*Front. Ecol. Environ.* doi:10.1890/080085 (2009)

A greater diversity of crops in a given area may reduce the amount of dissolved nitrogen compounds from fertilizers leaking into surrounding water bodies, where they wreak havoc with aquatic ecosystems.

Whitney Brouard, currently at the University of Louisiana at Lafayette, and Eugene Turner from Louisiana State University in Baton Rouge examined the relationship between various agricultural practices and levels of dissolved nitrogen compounds. They looked at data for 56 watersheds across the continental United States at both the beginning and the end of the twentieth century.

The duo discovered that about 45% of

the difference seen in the concentration of nitrogen compounds in the watersheds between 1997 and 2002 was due to variation in the biodiversity of crops on agricultural land in those areas.

## MICROMECHANICS

## Bacterial spin doctors

*Phys. Rev. Lett.* 102, 048104 (2009)

A micro-windmill driven by tumbling bacteria has been designed by Luca Angelani and his co-workers at the University of Rome La Sapienza.

Strikingly, the ‘bacterial wind’ isn’t blowing in any particular direction — a cogwheel rotor converts the uncoordinated impacts of many bacterial cells into directional rotation by means of the asymmetrical, sawtooth shape of its teeth. This makes the ‘bacterial motor’ akin to Brownian motors, long familiar in physics and biology, in which random motions are converted to directional ones by asymmetry in the environment.

But the self-propelled nature of bacteria makes this device subtly different: in effect, it converts chemical energy into mechanical work.

## EVOLUTION

## All for self

*Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0807679106 (2009)

Many plant species have accepted the trade-off between the benefits of greater self-reliance and the disadvantages inherent in inbreeding to switch from outcrossing to self-fertilization. However, little is known about the speed at which this happens, say

Stephen Wright of the University of Toronto in Canada and his colleagues.

By sequencing 39 nuclear genes from specimens of the selfing *Capsella rubella* and its outcrossing progenitor *Capsella grandiflora*, they estimate that the move in this species happened within the past 20,000 years. This is consistent with it occurring since the last glacial maximum, after which agriculture spread across Europe, producing a situation favourable to plants more capable of colonization — one of the advantages of selfing. Natural selection for guaranteed reproduction can thus lead to major changes and speciation over short periods of time, the authors speculate.

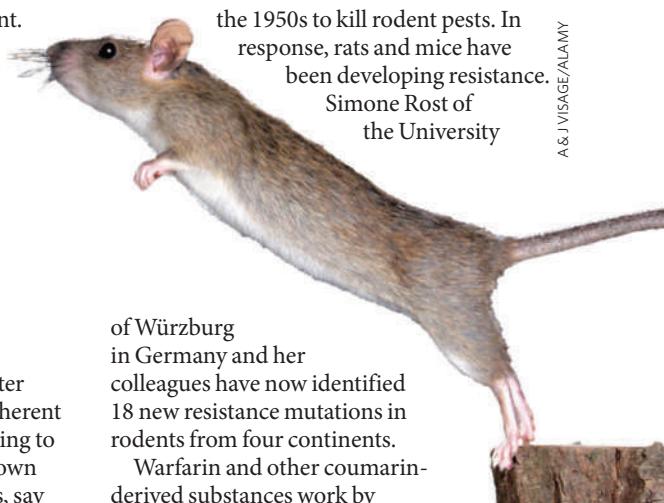
## GENETICS

## Rodent resistance

*BMC Genet.* doi:10.1186/1471-2156-10-4 (2009)

Pesticides derived from the compound coumarin have been used since the 1950s to kill rodent pests. In response, rats and mice have been developing resistance.

Simone Rost of the University



of Würzburg in Germany and her colleagues have now identified 18 new resistance mutations in rodents from four continents.

Warfarin and other coumarin-derived substances work by

## JOURNAL CLUB

**Nora Noffke**  
Old Dominion University,  
Norfolk, Virginia

**An astrobiologist considers life's oldest oxygen.**

The presence of atmospheric oxygen would have been necessary for the evolution of eukaryotes — organisms that group their genetic material into a membrane-bounded nucleus — so the question of when oxygen first became available is important in dating their rise. The availability of such oxygen is linked to the evolution of cyanobacteria, oxygen-producing microbes that appeared early in Earth's history and exist to this day.

Fossil microbial mats preserved in the Pongola Supergroup, a rock succession in South Africa, suggest that cyanobacteria were already highly diverse 2.9 billion years ago. But conclusive proof of their presence can be provided only by the presence of hydrocarbon biomarkers — stable chemical compounds found in the walls of single-celled organisms.

Work by Jacob Waldbauer at the Woods Hole Oceanographic Institution in Massachusetts and his colleagues focuses on biomarkers from shallow-marine deposits in the younger, 2.6-billion-year-old sedimentary rocks preserved in South Africa's Transvaal Supergroup. Detailed laboratory analyses extracted biomarkers called hopanes, possibly attributable to cyanobacteria, as well as steranes, biomolecules typically found in eukaryotes (J. R. Waldbauer *et al.* *Precamb. Res.* doi:10.1016/j.precamres.2008.10.011; 2008). The biosynthesis of steranes requires free oxygen; therefore, the fossil steranes imply that oxygen was readily available 2.6 billion years ago. This is at least 200 million years before a persistent oxygen-containing atmosphere is thought to have arisen.

Waldbauer *et al.* show that cyanobacteria had colonized the floor of Earth's ancient oceans by 2.6 billion years ago at the latest. Free oxygen has been available in the atmosphere ever since, and set the stage for the evolution of more complex organisms.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

repressing the enzyme VKOR, which is essential to blood coagulation. The mutations were in the gene *Vkorc1*, which codes for a key component of this enzyme, in the majority of 250 rodents trapped in areas where anti-coagulants are used.

### ZOOLOGY

## Restless rodents

*Ethology* 115, 217–226 (2009)

Burrowing mole-rats (pictured below) constantly dig new tunnels and change nests even when the soil is dry and tough. Subterranean rodents had been predicted to keep energy-demanding digging to a minimum during the dry season.

Radem Šumbera at the University of South Bohemia in České Budějovice, Czech Republic, and his colleagues used radio-tracking to follow the underground movements of ten silvery mole-rats (*Heliosciurus argenteocinereus*) in southern Malawi during a three-month period, then excavated and mapped the burrows.

slowing — or even preventing — it in others.

Darryn Waugh at Johns Hopkins University in Maryland, Baltimore, and his colleagues modelled climate and ozone interactions, both with and without ODS emissions, from 1960 to 2099. They found that ozone levels in the upper stratosphere recover decades earlier than expected owing to greenhouse-gas-induced cooling, which slows the chemical reactions that destroy ozone. By contrast, upwelling air masses hinder ozone formation and prevent recovery in the lower stratosphere of the tropics and southern mid-latitudes.

The authors say that these variations need to be taken into account when evaluating recovery of the ozone layer in the future.

### NETWORKS

## Know a good dentist?

*Phys. Rev. Lett.* 102, 058701 (2009)

If you need a dentist in London, who do you ask? Perhaps a friend who lives there? Even if they don't have a name, they can put you in touch with another friend who might. In this strategy, called 'greedy routing', you navigate the network of Londoners without knowing its global structure.

Now Marián Boguñá of the University of Barcelona in Spain and Dmitri Krioukov at the University of California, San Diego, prove that greedy is speedy. They show that greedy routing yields the fastest journey through networks such as the Internet. They suggest that switching to greedy routing could improve the Internet's speed.

### CANCER BIOLOGY

## Room to breathe

*Cell* doi:10.1016/j.cell.2009.01.020 (2009)

Improving the condition of tumour blood vessels may reduce the likelihood that the cancer will spread.

Abnormal blood vessels inside tumours impede the delivery of oxygen to cancerous cells as well as affecting the cells' sensitivity to chemotherapy. Meanwhile, oxygen-starved tumour cells are more likely to metastasize. To study this process, Peter Carmeliet of the Catholic University of Leuven in Belgium and his colleagues created mice that had only one functional copy of the *PHD2* gene, which encodes an oxygen-sensing protein called PHD2.

Tumour blood vessels in these mutant mice were not leaky like those in normal mice. And although tumours in the mutants grew normally, they received more oxygen and did not metastasize as often as tumours in normal mice.



J. ŠKUBA

The mole-rats switched nests about once a month, and on average each dug 0.7 metres of new tunnels daily and filled in 64% of old tunnels. Burrowing continued through the hot dry season, countering expectations that the rodents dig mostly while the soil is damp.

### ATMOSPHERIC SCIENCE

## Ozone in a warming world

*Geophys. Res. Lett.* doi:10.1029/2008GL036223 (2009)

The Montreal Protocol has been successful at phasing out ozone-depleting substances (ODSs), but scientists know that climate trends will also affect the ozone layer's recovery. New modelling work suggests this climate–ozone interaction could vary by region, hastening recovery in some areas while

# NEWS

## Kaputnik chaos could kill Hubble

Worst-ever orbital collision leads to calls for tighter regulation.

A cloud of debris spreading through low Earth orbit following the collision of two satellites poses a new risk to many scientific missions and may signal the demise of the Hubble Space Telescope. NASA is monitoring the increased threat carefully, and if it is as bad as some fear, the agency may have to cancel the proposed shuttle-servicing mission slated for later this year. Without that mission, the telescope's days are numbered, even if none of the new debris comes anywhere close to it.

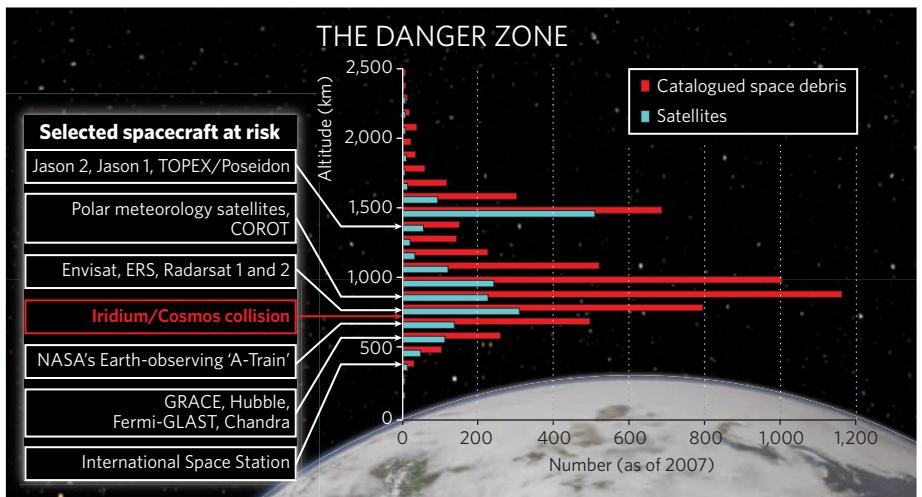
At 04:56 GMT on 10 February an active communications satellite owned by Iridium Satellite of Bethesda, Maryland, and a defunct Russian military-communications satellite collided some 800 kilometres above Siberia at more than 10 kilometres per second. The cloud of debris initially consisted of 600 objects large enough to be tracked by the US space-surveillance network, and experts expect that number to grow to more than 1,000 within the coming weeks. Simulations suggest there will be millions more pieces too small to track.

A preliminary analysis by researchers at the University of Southampton in the United Kingdom shows that a head-on collision between the satellites would have released some 50 kilojoules of energy per gram, about ten times the yield of TNT and perhaps a hundred times more than the energy released in China's 2007 test of an anti-satellite weapon. That test also broke up a satellite and exacerbated the debris problem, which has already affected some satellites (see 'Greatest hits'). "It's totally unprecedented," says Graham Swinerd, a reader in astronautics at Southampton. If the blow was only a glancing one, though — if, for example, the Iridium satellite was snagged by a boom believed to have been deployed on the Russian satellite, rather than hitting the main body — the situation may not be so bad.

Iridium, which operates a constellation of 66 low Earth orbit satellites providing satellite-phone services, says that it regularly monitors data about space debris but had no prior warning of the collision. There has been no official comment from the Russian government.

### Don't take the A-Train

The crash took place in a band of space heavily used by Earth-observing satellites, and space agencies are now closely monitoring the debris field as it spreads. NASA's 'A-Train' constellation and the Envisat mission of the European Space Agency (ESA) both orbit at very similar



altitudes to that of the crash, and are thus at particular risk (see graphic). "Right now we are doing a statistical analysis of what the increase in collision probability will be," says Heiner Klinkrad, who heads ESA's space-debris office in Darmstadt, Germany.

The increased risk to astronauts in the International Space Station seems to be "relatively low", according to Mark Matney, an orbital-debris specialist at the Johnson Space Center in Houston, Texas. But the collision puts in jeopardy a shuttle mission to repair the Hubble Space Telescope in May. The risk of impact for a mission to the space station is about 1 in 300, but for missions to the higher and more tilted

orbit of Hubble the risk is greater. Even before last week's collision, the added debris from the 2007 Chinese test had pushed up the Hubble mission's risk of a catastrophic impact to 1 in 185. NASA's usual limit on such risk is 1 in 200, so Matney describes the situation before last week as already being "uncomfortably close to unacceptable levels". "This is only going to add on to that," he says. Matney believes that the agency will know within a week or two whether the mission can go ahead.

### The blame game

In the immediate aftermath of the collision, experts were divided over whether Iridium should have seen it coming. "This should never have happened," says Geoffrey Forden, a space analyst at the Massachusetts Institute of Technology in Cambridge. Data provided by the US military showed that the two satellites would come within around half a kilometre of each other. "They should have manoeuvred Iridium away from it," he says.

But predicting satellite collisions is a tricky business, says Richard Crowther, head of the UK delegation to the UN Committee on the Peaceful Uses of Outer Space (UNCOPUOS). "It's actually a complex, time-consuming thing to do," he says. Even the best predictions are only probabilistic, and manoeuvring a spacecraft can involve costs and risks that outweigh the chance of collision.

Making avoidance easier is a matter of data and resources. Whereas Russia has its own system for tracking objects in space, the rest of the world depends more or less entirely on data

### GREATEST HITS

Space collisions to date

1991	Debris from one Russian satellite hits another, inactive Russian satellite.
1996	Cerise, a French satellite, is hit by debris from an Ariane rocket stage.
1997	Uncatalogued debris hits US NOAA 7 satellite (already inactive).
2002	Uncatalogued debris hits inactive Russian satellite.
2005	Debris from a Chinese rocket hits a US rocket.
2007	Uncatalogued debris hits the Meteosat-8 satellite.
2007	Uncatalogued debris hits NASA UARS satellite, already inactive.

Source: D. Wright, UCS



**HAVE YOUR SAY**  
Comment on any of our  
News stories, online.  
[www.nature.com/news](http://www.nature.com/news)

released by the US Department of Defense's Space Surveillance Network. The military keeps far better analyses of its data than those released to the public, according to Brian Weeden, a former analyst with the US Strategic Command, which oversees the Pentagon's sensor network. These analyses are used to look out for dangers to military and intelligence satellites, as well as some high-value civilian missions such as the International Space Station. Those data should be released in the interest of "the public good", says Jonathan McDowell, an astronomer at Harvard University who keeps track of satellite launches as a hobby. McDowell believes that nations should consider a multinational "space air-traffic control" to warn of such collisions.

More must also be done to prevent the further growth of orbital debris, adds David Wright, a researcher with the Union of Concerned Scientists, a non-profit group based in Cambridge, Massachusetts. At the moment, UNCOPUOS has a set of guidelines for limiting space debris, such as venting leftover propellant to prevent explosions. But Wright believes there should be binding rules for nations that launch into the cluttered region where the latest crash occurred. The debris created by each collision increases the risk of the next, and without action the number of incidents is bound to rise dramatically — possibly even leaving low Earth orbit completely unusable.

"I would like to see some sort of effort to make [existing guidelines] mandatory with some sort of enforcement mechanism," Wright says. Sanctions or fines should be imposed on negligent operators.

At the moment, the only legal framework for such action appears in the United Nations' 1967 Outer Space Treaty, which says that one nation can be liable for damage to another nation's satellite. But the wording of the treaty is too vague to be of much help in this case, according to McDowell. "Whose fault is it? Did one of them come up from behind?" he asks. "I don't think the road-traffic rules quite work." McDowell stops short of calling for an international treaty governing the low Earth orbit, but he does believe that "rules of the road" should be established to try and clear up the legal issues surrounding this and future collisions.

Liz DeCastro, a spokeswoman for Iridium, says she is unsure whether the company will pursue legal action against the Russian government.

**Geoff Brumfiel; additional reporting by Roberta Kwok.**

**See Editorial, page 936.**

## Tumours spark stem-cell review

A report claiming that unregulated transplants of human fetal neural stem cells led to tumours in a boy's brain and spinal cord is being hotly discussed by stem-cell researchers. Although the procedure took place well outside the scientific mainstream, it underscores the need for caution as clinical trials involving stem-cell transplantation move forward.

"This is the first documented case of tumours resulting from fetal-cell transplant," says Wise Young, a neurosurgeon at Rutgers University in Piscataway, New Jersey.

The paper, appearing in *PLoS Medicine*, describes a boy who received injections of fetal neural stem cells — for the neurodegenerative disease ataxia telangiectasia — at a Moscow clinic at the ages of 9, 10 and 12 (N. Amariglio *et al.* *PLoS Med.* 6, e1000029; 2009). When he was 13, researchers from the unrelated Sheba Medical Center in Tel Aviv, Israel, identified growths in his brain and spinal cord. The following year, spinal tumours were surgically removed. Within the growths, the team found cells that could not have arisen from the patient and must have come from at least two fetuses.

Gideon Rechavi, the team leader, says he has communicated with the Moscow clinic only through the boy's parents. He declines to disclose its identity, but says those involved apparently did not characterize the cells well. Rechavi adds that the boy's disease, which affects the immune system, may have made it easier for the tumour to grow.

During the past three months, three biotech companies have received regulatory authority to proceed with clinical trials that transplant fetal stem cells into the brain

or cells derived from embryonic stem cells into the spinal cord. All three — ReNeuron of Guildford, UK; Geron of Menlo Park, California; and StemCells of Palo Alto, California — have presented preclinical data on safety and efficacy for their cells.

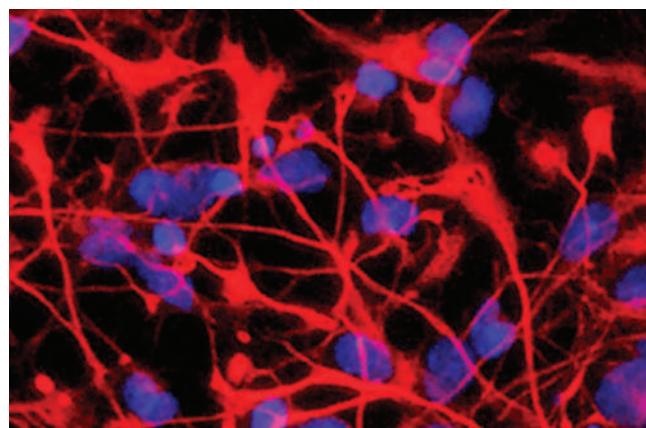
John Sinden, ReNeuron's scientific co-founder, says that "there are almost no similarities" between the Moscow procedure and his company's clinical trial. He argues that one isolated case doesn't threaten the whole field.

It might, though, serve as a cautionary tale, says Steven Goldman, chair of neurology at the University of Rochester in New York and an adviser to the US Food and Drug Administration. He believes that many techniques that induce stem cells to expand in culture render them more likely to grow inappropriately. "I hope it makes people think twice about how cells are isolated and maintained before transplant," he says.

"This paper is actually a really good thing for the field," says Insoo Hyun, a bioethicist at Case Western Reserve University in Cleveland, Ohio, who has led efforts to help patients assess the risks of stem-cell procedures. "People who undergo these cell-based therapies have a responsibility to share any adverse events that happen," he adds.

However, Theo Palmer, a neurosurgeon at Stanford University School of Medicine in California, says the new report is unlikely to do much to help researchers understand or avoid risk. The Moscow cells may or may not have qualified under the strict regulations for human stem-cell trials. "This makes it impossible to draw parallels to other neural-stem-cell studies performed under rigid safety criteria," he says.

Monya Baker



Neural stem cells in culture: use them carefully.

# Medical research scores big in US stimulus bill

Congress votes to give generous boosts to other agencies too.

The National Institutes of Health (NIH) could call him Saint Specter. Cancer survivor Arlen Specter of Pennsylvania was one of three moderate Republican senators whose votes were crucial in pushing a \$787-billion economic stimulus package through the US Congress. In the Senate, Specter had introduced an amendment to triple the NIH's share of the spending, calling it "scandalous" that the country doesn't spend more on medical research. His proposed \$10 billion of extra spending for the NIH over two years made it through to the final version of the bill agreed by both houses of Congress on 13 February.

That spending on biosciences did well compared with other science spending is not a great surprise. "Generally, the NIH gets all the big funding because people think they're going to save lives that way," says Vernon Ehlers, a physicist and member of the House of Representatives (Republican, Michigan). "Members of Congress feel a vote for the NIH is a vote for good health." What has been a surprise to many is the sheer amount of beneficence. Although there was plenty of money for physical sciences, too, in the bill, the sums were nothing like so spectacular.

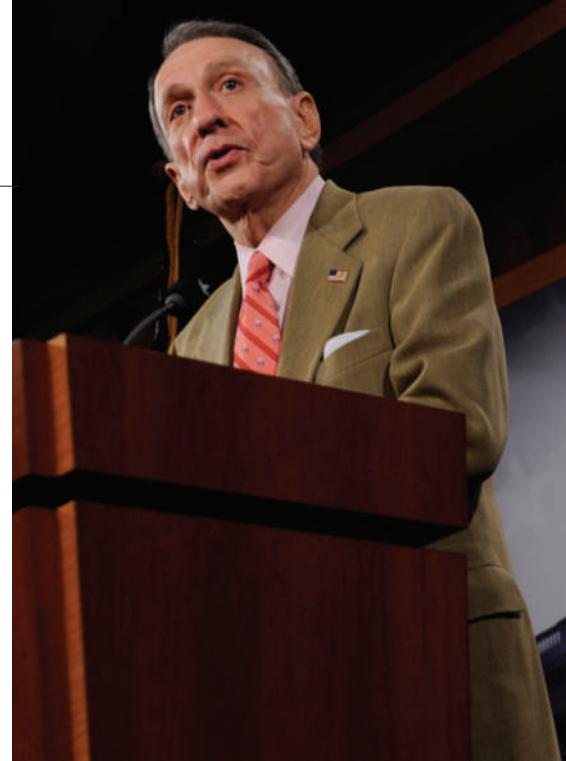
The stimulus package, which is one-third tax cuts and two-thirds spending, contains

some \$21.5 billion for scientific research and development across all agencies, according to an analysis by the American Association for the Advancement of Science. In relative terms, the National Science Foundation (NSF) did even better than the NIH, as its \$3-billion award is almost half its annual budget (see 'The stimulus bill: who got what'). The Department of Energy, while receiving a relatively modest \$1.6 billion for its research-oriented Office of Science, will get a huge sum agency-wide — about \$40 billion for a variety of energy and energy-efficiency projects.

"The stimulus package is a singular event in the history of science funding," says John Marburger, former presidential science adviser and head of the Office of Science and Technology Policy under George W. Bush.

President Barack Obama was expected to sign the bill into law shortly after *Nature* went to press. After that, agencies have 60 days to present spending plans to the White House, and then cheques are meant to be cashed quickly: most spending must be completed by 30 September 2010.

The bill is meant to get government money into the faltering economy speedily. The speed with which it was put together meant that individual members of Congress, such as Specter,



Senator Arlen Specter's efforts got the National Institutes of Health \$10 billion in extra money.

S. WALSH/AP

were able to wield huge influence. A version of the bill that circulated as it was being finalized was covered in strike-throughs and hand-scratched marginalia that exemplified how quickly hundreds of millions of dollars in science funding could appear, or disappear, at the stroke of a pen. Marburger, for one, questions whether the Obama administration, which is still naming the heads of key science agencies (see story below), was able to exert as much control over the legislation as it wanted.

What began as a bill that was supposed to be limited to scientific infrastructure ended up supporting mostly research and grants, notes Michael Lubell, director of public

## John Holdren: adviser on science, fish and wine

**John Holdren, US President Barack Obama's choice as chief science adviser, faced a single criticism during his otherwise gentle Senate confirmation hearing on 12 February: over the years he had warned of disasters that have yet to materialize.**

Given Holdren's qualifications as a physicist and engineer with nearly four decades of work on energy, climate and nuclear proliferation behind him, it was about all the critics had to go on. The job fell to Louisiana Republican David Vitter, who peppered Holdren with gloomy assessments dating back to 1971 about overpopulation, an unspecified

eco-catastrophe, thermonuclear war and global warming.

Calmly pointing out that he was 26 in 1971, Holdren said one of the things he has learned since is that predicting the future is "difficult". And although he said he no longer thinks it useful to focus on population, he didn't yield any ground on global warming: "I think it is important to call attention to the dangers that society faces," he told Vitter.

The Senate is expected to confirm Holdren's nomination soon, along with that of Jane Lubchenco, a marine biologist from Oregon State University who will head the National Oceanic and Atmospheric

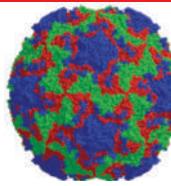
Administration (for coverage of their confirmation hearings, see <http://tinyurl.com/dyljpt>).

As chief of the Office of Science and Technology Policy (OSTP), Holdren will lead a staff of about 60 that provides technical advice on daily decisions, budgets and the full range of federal policy. In many ways, friends and colleagues say, he has been training for this position his entire career.

"John is enormously focused and incredibly persuasive," says Henry Lee, who until recently co-taught a course with him on energy technology, policy and economics at Harvard's Kennedy School of

Government. "I think in a matter of months he will be a major intellectual driver in this administration."

Holdren has a history of getting things done. Beginning in 1973, he built the University of California Berkeley's fledgling Energy and Resources Group, the first interdisciplinary graduate programme on the campus, into a major initiative that continues to produce leading research on everything from state and local air pollution regulations to international energy and climate policy. "It is one of the only interdisciplinary energy centres that has not only persisted since



## THE GENOMICS OF THE SNIFLES

Cold-virus sequences may be the key to new therapies.  
[www.nature.com/news](http://www.nature.com/news)

J.-Y. Sgro/UW MADISON



affairs for the American Physical Society in Washington DC. Many hope that such broad cash infusions mark the start of sustained future investments in basic science funding in a way that one-off spending on infrastructure would not. "What was put in the stimulus is what was supposed to be in appropriations from last year," says Ehlers, referring to the attempts to boost physical sciences through the America COMPETES Act. "I hope it's part of a sustained effort," he adds.

Congress has not yet passed final spending bills for fiscal year 2009, although it is expected to work on an omnibus spending bill in the week beginning 23 February. Around then, the Obama administration is expected to release its first sketch of what it wants in its 2010 budget.

then, but literally grown into a major hub of action on energy and climate and sustainability," says Dan Kammen, who studied there and assumed Holdren's position after he left for Harvard.

In 1996, Holdren came to the Kennedy School to head the Belfer Center's Science, Technology and Public Policy Program; he expanded its focus on energy and climate policy while bolstering research into nuclear non-proliferation issues. Centre director Graham Allison says that Holdren was an ideal candidate for the job, but he is only half joking when he says it was Holdren's love of fishing that made the deal possible. While still at Berkeley, Holdren had bought a home near Woods Hole, Massachusetts, with

easy access to top-notch saltwater fly-fishing. Allison discovered he was looking to make a move, contacted him and eventually sweetened the Harvard deal by offering to take Holdren and his wife Cheryl on a fishing trip in the Caribbean (paid for by Allison).

Friends say Holdren deploys the tools of science in his life as a whole. His fishing database, where he keeps track of key variables such as time, location, tide and water temperature, along with records of which fish are biting on what, is legendary. Holdren has also developed an algorithm for comparing wine ratings to price lists, and he sends a list of the best-buys to friends each Christmas. "But you have to act on it quickly, because sometimes it moves local

markets," Allison says.

Holdren typically took the bus from Woods Hole — where he was director of the Woods Hole Research Center — to Cambridge for work each Tuesday and stayed the night with Dan Schrag, a geochemist in Harvard's earth and planetary sciences department. Holdren would then attend the department's regular Wednesday breakfast meeting, a freewheeling session for professors, postdocs and guests, taking the bus back on Wednesday night.

Last week's breakfast began with Holdren's appointment before shifting to the economic crisis, science as economic stimulus (see story above) and how a severe recession might affect universities. Schrag says it might be a while

before anybody knows just how effective Holdren is in his new position, but it bodes well that he has already secured four deputies (the last administration cut the number down to two), restored the OSTP's place on the security council and negotiated workspace in the Eisenhower Executive Office Building next door to the White House.

"I've gotten e-mails from John saying, 'Can you get me a list of X in the next two hours?'. That's a good sign," he says, as presidential advisers need to influence decisions in the White House on the order of days or even hours. "The question is whether he and his deputies will be in that inner circle that is making decisions in the West Wing."

Jeff Tollefson

## THE STIMULUS BILL: WHO GOT WHAT

### National Institutes of Health

**Stimulus:** \$10 billion  
**2008 budget:** \$29.6 billion  
**Highlights:** \$7.4 billion will go to the institutes and centres. \$800 million will remain in the director's office, of which half may be used for initiatives under the Road Map for Medical Research, which focuses on cross-disciplinary, highly innovative and translational research. \$1.3 billion will be spent on construction, repairs and shared instrumentation at non-federal research facilities. \$500 million will fund high-priority repair, construction and improvement projects on the agency's campus in Bethesda, Maryland.

### National Science Foundation

**Stimulus:** \$3 billion  
**2008 budget:** \$6.1 billion  
**Highlights:** \$2.5 billion will go towards external research

grants, including \$300 million for instrumentation. A separate allowance of \$400 million will go to construction of major facilities.

### Department of Energy

**Stimulus:** About \$40 billion  
**2008 budget:** \$23.9 billion  
**Highlights:** Includes \$11 billion for the electric grid, \$5 billion for weatherproofing homes, \$3.4 billion for fossil energy R&D and \$2 billion for battery research. The Office of Science, which funds basic research, receives \$1.6 billion. A separate \$400 million will kick-start the Advanced Research Projects Agency-Energy.

### NASA

**Stimulus:** \$1 billion  
**2008 budget:** \$17.2 billion  
**Highlights:** \$400 million is for science, to accelerate Earth science climate-research missions and increase

supercomputing capabilities. Another \$400 million could be spent on rocket development to shrink a 'gap' in human spaceflight capability caused by retirement of the space shuttle.

### National Oceanic and Atmospheric Administration

**Stimulus:** \$830 million  
**2008 budget:** \$3.9 billion  
**Highlights:** Directs \$170 million to fill critical gaps in climate modelling and to establish climate data records.

### National Institute of Standards and Technology

**Stimulus:** \$580 million  
**2008 budget:** \$737 million  
**Highlights:** \$360 million would go to construction and renovation of facilities, with \$180 million for funding competitive external applications for research science buildings.

Here's a look at some of the big winners in the stimulus bill.

**National Institutes of Health:** \$7.4 billion divided among the agency's scientific institutes and centres will fund grants from a backlog of 14,000 investigator-initiated 'R01' grants already reviewed and categorized as "highly meritorious". It will also fund new R01 applications for projects that could reasonably make progress in two years. The agency will add

supplemental funding to existing grants and fund new "challenge grants" aimed at thorny problems.

"We are confident that we can spend the funds that Congress has allocated both responsibly and quickly," says Raynard Kington, acting NIH director.

Advocates of biomedical research, who have lobbied fruitlessly for substantial NIH increases in recent years, were ecstatic. But they



**GOT A NEWS TIP?**  
Send any article ideas for  
Nature's News section to  
[newstips@nature.com](mailto:newstips@nature.com)

I. GOOD  
K. CAMPBELL/GETTY

► also cautioned against repeating history, in which a budget doubling between 1998 and 2003 was followed by five years of flat funding, leading to a glut of investigators and plummeting application success rates. "We're hopeful that this represents a first step towards sustained growth," says David Moore, senior director of government relations at the Association of American Medical Colleges in Washington DC.

Kington notes: "We are being very careful to focus on funding that only covers the two years of the stimulus package. There will be relatively little, if any, money that entails a four-year commitment."

**National Science Foundation:** The \$3-billion stimulus award is a vast amount for an agency with an annual budget of only twice that. Will the NSF have trouble distributing the money quickly and wisely? "It's not a problem at all," says Rita Colwell, former NSF director, who believes that the agency could use up all of the stimulus money with existing grant applications. Michael Turner, former assistant director for mathematical and physical sciences at the NSF, says that grant-application success rates have been very low, around 20%. By increasing that rate to just 30–40%, the agency could use up the \$2.5 billion that has been designated for research. "They have many more proposals that are highly worthy," he says. Another \$400 million has been designated for the construction or rehabilitation of facilities.

**Department of Energy:** Lubell says that the Office of Science should have no trouble spending its \$1.6-billion award even without funding grants. With the help of the department's national laboratories, he drew up a list of infrastructure needs alone that came to that figure. In a statement, acting Office of Science director Pat Dehmer says her office is "putting the mechanisms in place to get the funding out quickly with strong transparency and accountability controls".

The bigger question is how well the rest of the agency will fare. Energy secretary Steven Chu has been given sums totalling \$40 billion, that dwarf the accustomed budgets in the parts of the agency targeted. For instance, spending on fossil-fuel energy research and development was just under \$750 million last year; the stimulus bill will dump \$3.4 billion in that arena.

Much of the energy department's new money can be distributed in big chunks, because many energy demonstration projects require loan guarantees worth hundreds of millions of dollars. ■

**Eric Hand and Meredith Wadman**



Ancient silk strands have been found in artefacts from Pakistan's Indus valley.

## Rethinking silk's origins

Did the Indian subcontinent start spinning without Chinese know-how?

New findings suggest that silk making was not an exclusively Chinese technological innovation, but instead arose independently on the Indian subcontinent.

Ornaments from the Indus valley in east Pakistan, where the Harappan culture flourished more than 4,000 years ago, seem to contain silk spun by silk moths native to the region. What's more, the silk seems to have been processed in a way previously thought to have been a closely guarded secret within China.

There is hard and fast evidence for silk production in China back to around 2570 BC; the newly discovered objects are believed to date from between 2450 BC and 2000 BC, making them similarly ancient. There have been no previous finds of manufactured silk at sites outside China before about 1500 BC.

"This is the first evidence for silk anywhere out of China at such an early date," says Irene Good of Harvard University in Cambridge, Massachusetts, one of the authors of the study. "It was a complete surprise."

The objects come from two sites in the Indus valley: the city of Harappa itself, the hub of the Indus civilization, and Chanhudaro in Sindh province, about 500 kilometres to the south. They were collected from archaeological excavations in 1999 and 2000 conducted by the Harappa Archaeological Research Project (HARP), a US–Pakistan collaboration. Because of the sheer volume of artefacts amassed so far, they have only recently been studied in detail.

Good, working with HARP directors Richard Meadow of Harvard University and Jonathan Kenoyer of the University of Wisconsin in Madison, used an electron microscope to look at the fine structure of silk strands found in necklaces and bangles.

The precise shape of the individual silk

threads — determined by the shape of the orifice through which they are extruded — is characteristic of the species of silk moth that produced the strands.

In a paper in the journal *Archaeometry*, the researchers show that the Harappa samples — two metal ornaments — contain silk from species of *Antheraea* moths indigenous to south Asia (I. L. Good, J. M. Kenoyer and R. H. Meadow *Archaeometry* advance online publication doi:10.1111/j.1475-4754.2008.00454.x; 2009). The origin of the Chanhudaro silk, threaded through soapstone beads, is less clear, but it may be from one of the same species. Chinese silk comes from the domesticated silk moth *Bombyx mori*.

The Harappan silks seem to have been made by a process called reeling, in which the strands are collected on a bobbin rather than being twisted in short segments into a thread. The researchers say that reeling was thought to have been part of a silk technology known only to China until the early centuries AD. Now it seems that knowledge was not so exclusive.

"Archaeology in early China is showing increasingly there were connections outside China," says Shelagh Vainker, a silk expert at the Ashmolean Museum in Oxford, UK. "It doesn't seem unreasonable." But she sees evidence for silk production in China "significantly earlier" than 2500–2000 BC, which would suggest China could still claim priority.

"I believe that the people of the Indus civilization either harvested silkworm cocoons or traded with people who did, and that they knew a considerable amount about silk," says Good. She does, however, acknowledge that some might find this challenging: "National pride is at stake with such a discovery as this." ■

**Philip Ball**



Austria is one of a handful of nations that have banned cultivation of genetically modified corn.

# European disarray on transgenic crops

Forthcoming decisions set to bring disagreements to a head.

Europe has failed to reach a decision on whether France and Greece should be ordered to lift their national bans on cultivation of a genetically modified maize (corn) known as MON801, the only GM crop so far approved for use at the European level. The failure of France, Greece, Austria and Hungary to adhere to European Union (EU) legislation threatens a political crisis in the EU, which is under pressure from the World Trade Organization to live up to its laws in this matter.

The European Food Safety Authority (EFSA) in Parma, Italy, which is responsible for scientific risk assessments on GM crops in Europe, has reported that there is no case against MON801, which has been modified to be resistant to insects. But on 16 February, a committee of experts appointed by each of the 27 EU countries was unable to muster the 'qualified majority' — one that

represents 62% or more of the Union's population — needed to decide whether to lift the bans. The result echoes a similar impasse last December, when the committee failed to reach the qualified majority needed to order Austria and Hungary to lift their bans on the crop.

The World Trade Organization claims that the EU is out of line with its own legislation, and that the national bans are effectively denying countries that use biotechnology in agriculture the access to the European market they are allowed under trade rules. A series of looming decisions on these issues is set to bring the matter to a head over the next month.

Most of Europe's citizens, and many of its governments, have long been hostile to agricultural biotechnology. A number of member states did not implement the EU's decision to approve MON801 in 1998, and this led the

**"The legal framework is science-based, yet getting a decision is very difficult."**

European Commission, the EU's executive, to develop a new legal framework that ensured the tightest possible safety controls based on transparent risk-assessment analysis.

## Dissension in the ranks

That framework, which finally came into force in 2004, allows the EU to approve GM crops when there is no scientific evidence of danger to health or environment; once approved the crop may be cultivated in all 27 member states. It also includes provision for revisiting an approval if new scientific evidence of possible danger should come to light. Seven countries in the EU now grow MON801, and in most others farmers are free to do so should they so wish. But Austria, Greece and Hungary have consistently refused to allow cultivation, claiming that the risk assessment needs to be updated.

Last year, France, which had been one of the EU's largest users of the technology, joined the

**DRUG BANISHES BAD MEMORIES**

Blood-pressure pill quells fear.

[www.nature.com/news](http://www.nature.com/news)

FOCUS FEATURES/KOBAL COLLECTION/D. LEE

refuseniks. It sent the EFSA a thick dossier of new data that it argued raised safety concerns. But the EFSA concluded that none of the supposedly new scientific evidence provided by these four countries "would invalidate the previous risk assessments of maize MON810".

Given the committee's failure to make a decision, the dossier on France and Greece will now have to be considered by the environment ministers of the EU nations within the next three months. The ministers will also consider the cases of Austria and Hungary when meeting in council on 2 March.

If the council of ministers fails to reach a qualified majority in favour or against, the commission itself has to make the decision. "Ad infinitum, ad absurdum," says an insider at the environment ministry. "The legal framework is science-based, yet getting a decision is very difficult."

**Lack of trust**

Natalie Moll of EuropaBio, the biotech industry lobby group in Brussels, says that the regulatory committee's failure to bring an end to cultivation bans "gives the usual European mixed signal — that we have the toughest scientific approval system in the world, but we don't trust it".

In the meantime, the EFSA is working through its backlog of applications for the cultivation of 13 other GM crops. It has already approved two more insect-resistant maizes — BT1507, which is jointly owned by Pioneer and Dow Agrosciences, and BT11, owned by Syngenta. Both approvals will be discussed by the regulatory committee on 25 February.

As the timetable for key decisions tightens, many politicians are entrenching their positions. After a risk-analysis report by the French food-safety committee giving MON810 a clean bill of health was leaked to the press last week — it had previously been suppressed — prime minister François Fillon rushed to Brussels to insist that France would maintain its ban whatever the EU decided. At the same time the state government of Bavaria stopped all field experiments on GM crops, confining them to greenhouses, and the Bavarian environment minister Markus Söder announced his intention to block cultivation of GM crops in the state.

Meanwhile, public opposition to GM crops may be slipping. In a Eurobarometer public-opinion survey published last year, the percentage of those who said they were against GM crops fell from 70% to 58%. ■

**Alison Abbott**

# MRI modified for better images

A simple change to magnetic resonance imaging (MRI) machines will provide more uniform coverage at higher powers as well as more room for portly patients. In a market set to be worth more than \$5 billion by 2010, the new technology may offer an easier way to get to the high-field machines manufacturers and clinicians see as the next target for hospital imaging.

MRI machines use a magnetic field to get hydrogen atoms in the body spinning in a particular way, then knock them off-balance with a radio wave. The small radio-frequency signals given off by the recovering nuclei provide the imaging data. In their new version of the technology, Klaas Pruessmann at the University of Zurich, Switzerland, his student David Brunner and their colleagues removed the radio-frequency coil used to tumble the nuclei from an MRI machine built by Philips Healthcare and replaced it with a system that could do the same job from up to 5 metres away. The university has filed for patents on the technology, which is described on page 994 of this issue.

"It's a completely new approach to exciting the signal in MRI," says Andrew Blamire, an MRI expert at the nuclear magnetic resonance centre in Newcastle, UK.

"Claustrophobia is a widespread problem in clinical MRI," says Pruessmann. Removing the coil from the machine provides a less constraining cavity. But the potential advantages go further than the patient experience. The easily made change of approach may allow designers of increasingly powerful MRI machines to overcome some of the technical hurdles that trouble them.

Standard clinical MRI scanners use magnets with field strengths of about 1.5 Tesla and radio-frequency signals of about 64 megahertz, but the fastest growing part of the market at the moment is for higher power, 3-Tesla machines. Even more powerful magnets would be better; they magnetize more nuclei, giving stronger signals. But stronger fields need higher frequency, shorter wavelength radio-frequency signals, and wavelengths shorter than the size of the subject — a human head, for example — lead to some areas producing either a zero signal (blind spots) or a very strong signal (hot spots). Magnets with strengths of 7 Tesla can produce super-clear images in some places, but because the corresponding radio-frequency wavelength is just 12 centimetres in human tissue, the

results are dark and undecipherable in others.

Pruessmann's technique uses a travelling wave rather than a standing wave, with the cylindrical conducting tube lining the machine functioning as a waveguide for a signal transmitted by an outside antenna. The approach could, in principle, allow designers to avoid many of the effects of hot spots and blind spots.

The technique was unveiled last May at the International Society for Magnetic Resonance in Medicine meeting in Toronto, Canada. This provoked Graham Wiggins of the Center for Biomedical Imaging at New York University's Medical Center to build his own version. He found that in whole body scans, the inhomogeneity wasn't completely removed, but the positions of the blind spots and hot spots were different. "I've spent so long trying to change that pattern and couldn't," says Wiggins. He can now see areas of the brain that were previously dark to MRI. "That's exciting," says Wiggins, who is working with one of the big industrial MRI players — the health-care arm of Siemens, which is based in Erlangen, Germany.

"Most of the major manufacturers have a significant research interest in clinical imaging above 3 Tesla," says Matthew Clemence, a senior scientist with Philips Healthcare based in Reigate, UK — one of those manufacturers. GE Healthcare, headquartered in Chalfont St Giles, UK, is another. "The biggest benefit of this technology is in reducing the image shading that occurs when standing waves are generated in the subject," says Jason Polzin, chief engineer of global magnetic resonance at GE. This is one of a variety of solutions to this problem, Polzin adds. Another approach is to use multiple radio-frequency coils within the scanner, which is "technologically fascinating but insanely complicated", according to Wiggins, and adds even more bulk to the internal space.

Pruessmann, although aware of the commercial interest for clinical MRI, sees wider applications for the new technique. "This introduces a whole new way for thinking to this field," he says, pointing to possibilities such as screening large numbers of biological samples or laboratory animals all at once. "Giving people a new degree of freedom will hopefully lead to things that we haven't thought of at all." ■

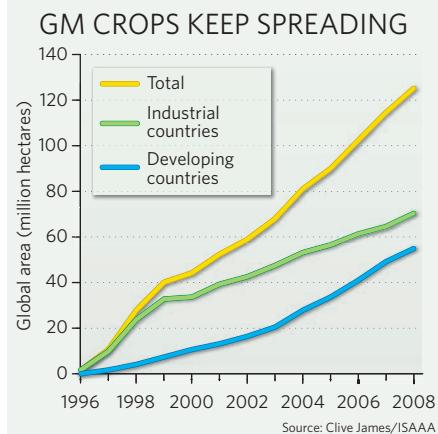
Katharine Sanderson

## Developing countries boost spread of GM crops

Planting of genetically modified (GM) crops grew by 9.4% in 2008 to 125 million hectares worldwide, achieving a total value of US\$7.5 billion, according to a report published on 11 February by the International Service for the Acquisition of Agri-biotech Applications, a non-profit biotech industry group.

In total, 15 developing countries and 10 industrial countries planted GM crops last year (see chart). The United States, with 62.5 million hectares, planted half of the total, the figures show.

Three countries grew GM crops commercially for the first time, including two in Africa. Egypt opted for biotech maize (corn) and Burkina Faso for cotton. Together with South Africa, which has grown biotech cotton, maize and soya beans since 1998, they bring the continent's total of GM-growing nations to three. Bolivia, 2008's third GM adopter, grew soya beans.



## University reforms stall as French protests surge

French scientists voted on 12 February to continue a national strike that began on 2 February in protest at proposed government reforms of academic careers (see *Nature* 457, 640–641; 2009). Valérie Pécresse, France's science minister, had offered a two-month mediation to revise the draft decree implementing the changes, but to no avail.

The row gathered momentum with nationwide street demonstrations on 10 February; Paris alone saw as many 40,000 researchers and students turn out to protest.

Researchers have now expanded their demands to include reversing the proposed transformation of the CNRS, Europe's largest basic-science agency, into a research council, and other reforms of the university and research systems.

The government lost further support as university presidents, top scientists

and a slew of research bodies called for the reforms to be rethought in closer consultation with the research community. For a longer version of this story, see <http://tinyurl.com/a9roy5>

## Budget crisis drives Harvard to redundancies

Harvard University will offer voluntary redundancy to 1,600 or so of its non-faculty employees. The offers are to target staff who are 55 years or older and have worked for the university for at least a decade.

The "early retirement program" was unveiled on 11 February and is the latest sign of the budget crisis facing the world's wealthiest university. The university's endowment, which in June 2008 stood at US\$36.9 billion, shrank by 22% late last year (see *Nature* 457, 11–12; 2009). The university estimates that the drop may reach 30% by the end of its fiscal year in June.

To cope with the crisis, departments have frozen faculty hiring and have begun to draw up contingency plans in case large building projects — including a proposed new science complex — are put on hold.

## Safety precautions delay start-up of hadron collider

The Large Hadron Collider (LHC), the world's most powerful particle accelerator, will see no collisions until late October, more than a year after its planned start date.

Officials from CERN, the LHC's host laboratory near Geneva, Switzerland, had hoped to restart the machine by summer 2009, after it was seriously damaged during power tests last September. But on 9 February CERN announced a further delay, citing additional safety protocols and complex repair schedules as the reasons.

With a short technical stop over Christmas, the LHC will run through to autumn 2010, but it will accelerate its protons to just 5 teraelectronvolts (TeV),

## SOUNDBITES FROM CHICAGO

Some insights from this year's meeting of the American Association for the Advancement of Science.

**"We are basically looking now at a future climate that is beyond anything we've considered seriously in climate model simulations."**

Chris Field co-chair of the Intergovernmental Panel on Climate Change's second working group

**"[Cloning a Neanderthal] starting from the DNA extracted from a fossil is and will remain impossible as far as I can see into the future."**

Chris McCarthy AAAS president



A superconducting magnet is removed at the LHC.

well below its designed 7 TeV. Pauline Gagnon, an experimental physicist with ATLAS, one of the LHC's detectors, thinks that at the reduced power "the chances of finding something will be limited", but that the LHC might still see some new physics.

For a longer version of this story, see <http://tinyurl.com/cy29f5>.

## GSK backs patent pool for neglected diseases

Pharmaceutical giant GlaxoSmithKline (GSK) will share the patented knowledge it uses to develop medicines for neglected tropical diseases (NTDs) — including malaria and tuberculosis — with other drugs companies, governments and non-governmental organizations.

Andrew Witty, GSK's chief executive, proposed the voluntary 'patent pool' in a 12 February speech, and called on other drugs firms to open up access to intellectual property relevant to NTDs.

Starting this year, Witty announced, GSK will also cap its prices for patented medicines in poor countries at 25% of what it charges in developed countries. It will also reinvest 20% of the profit it makes from selling medicines in poor nations into health-care infrastructure projects in those countries.



# THE GENOME OF THE AMERICAN WEST

What does it mean to save a species? For some, preserving the American bison means keeping its genome pure, finds **Emma Marris**.

James Derr has just eaten a large grass-fed-bison steak topped with onions, the banquet dinner at a meeting of the American Bison Society in Rapid City, South Dakota. As he sips his wine, conservationists and managers of nature reserves — some wearing Stetsons and some wearing polar fleeces — approach to pay their respects to the man who could be considered the godfather of the bison genome.

Derr, a geneticist, is trying to reverse more than a century of human interference with the American bison (*Bison bison*). Those shaggy symbols of the American West were driven to the brink of extinction in the last half of the nineteenth century and then saved on ranches and in zoos. But the conservation efforts came at a cost. Most of the bison alive today have cattle genes rattling around in their genomes, left over from early efforts to interbreed the two species.

Derr has almost single-handedly started a movement to preserve the original bison, complete with its unadulterated genome. He has managed to persuade everyone from federal officials to private conservation leaders that they should care about the cattle genes hiding in bison. And he is convinced that his

approach — managing the genome rather than the animals — could be a model for conserving other large mammals.

Wildlife managers have considered the genetic diversity of animals for some time, and animals in captivity have often been bred to preserve genetic diversity. But those were blunt approaches. Now, armed with genomic tools, researchers are starting to look at specific sequences in the genome, and are raising questions about what the fundamental unit of conservation should be. Most people see preserving wildlife as a matter of saving individuals; if all the individuals die out, the species becomes extinct. But that reasoning looks simplistic when considered at the genomic level. If the genes of a species change enough — through interbreeding, for example — that species will cease to exist even if individuals that look something like the original continue to thrive.

Although some species interbreed naturally, humans have forced other mix-ups, and those are the cases that most worry Derr. "Species conservation is more than skin deep," he says. "It is more than how they look, it is how they are — that's the genome."

This purist approach has won converts,

but other conservationists say that the bison is an exception, and that for many species, the situation is so dire that they don't have the luxury of worrying about a few stray genes from related species.

## Where the buffalo roamed

The troubles facing bison are relatively recent. When explorers Meriwether Lewis and William Clark voyaged across the continent at the beginning of the nineteenth century, tens of millions of bison roamed in massive herds from the Mississippi to the Rockies, and from Mexico to the Arctic. By the end of the century bison numbers had dropped to a few hundred, courtesy of long-range rifles and professional hunters.

At this point, a few forward-thinking people decided that the bison should not be driven to extinction. Charles Goodnight, the famous Texas cattle rancher, and a handful of other people made room for the animal on their private ranches. In 1905, then US President Theodore Roosevelt and William Hornaday, head of the New York Zoological Society (now known as the Wildlife Conservation Society), founded the American Bison Society, which collected bison and established herds in a few reserves



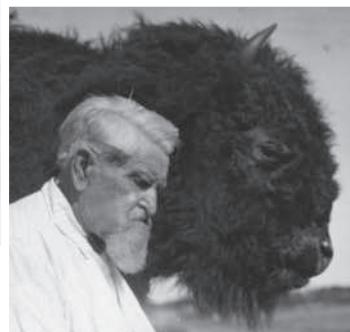
Charles Goodnight (right) was a pioneer of bison conservation.

PANHANDLE-PLAINS HISTORICAL MUSEUM, CANYON, TEXAS

in Montana, Oklahoma and South Dakota. A small herd, perhaps 30 in number, was still roaming Yellowstone National Park. According to Derr, all the bison in the United States today — there are now up to a million of them, mostly on private ranches — can probably be traced back to fewer than 200 bison.

Half a century after those conservation efforts started, Derr saw his first bison. He was a boy, growing up in Enid, Oklahoma, just north of a small town named Bison. His father raised cattle, and Derr lived mostly outdoors, hunting and fishing. The bison that young Derr encountered on local ranches did not act like his father's cattle, and they left a lasting impression. "They are much more athletic, very smart and incredibly quick and strong," he says.

In college, Derr studied biology with an eye to being a wildlife manager. But he soon decided "that I could have a lot bigger impact if I worked in the lab". He switched to genetics and is now a professor at Texas A&M University in College Station. In the 1990s, he put together a proposal to study how the bison had passed through that 200-individual bottleneck with seemingly no ill effects. Unlike many other species that have bounced back from small numbers, bison have



James Derr is pushing to conserve bison at a fundamental level, the genome.

retained a decent level of genetic diversity.

"It is a pretty darn amazing story," says Derr. At their fewest, bison were preserved by people sprinkled throughout the continent. "Just by luck, it seems that they were able to preserve, with the few bison that they got, much of the genetic diversity that was in those areas."

Another reason the bison came so easily through the bottleneck, he says, might be because they had already squeezed through earlier genetic restrictions. In these previous population crashes, the inbreeding required to re-expand the population exposed deleterious recessive genes, and those animals died, in effect purging the genome of lethal traits.

But while doing this work, Derr found foreign genes in the bison genome. Goodnight and the other ranchers may have saved the bison, but they also experimented with them. Cattle (*Bos taurus*) and bison are not too distantly related to produce fertile offspring — they diverged between 1 million and 1.5 million years ago but have the same number of chromosomes, with many genes in the same order. Although the two animals are not known to mate voluntarily, ranchers forced a few crosses in the early twentieth century, leaving the fingerprints of cattle on the nuclear and mitochondrial DNA of most bison alive today. Derr has found cattle genes in eight of fourteen public herds in the United States and Canada, and in all but one of the 50 private herds he has looked at (N. D. Halbert and J. N. Derr *J. Heredity* **98**, 1–12; 2007).

#### If it looks and acts like a bison ...

When he first started his work, Derr was worried "that there might not be any uncontaminated bison herds". But why should he care? These bison, apart from a few very strange specimens in Texas with white faces,

look like bison. They seem to act like bison. For conservationists who are interested in restoring the relationships that once prevailed in prairie ecosystems, that is probably good enough.

But for those interested in 'saving the bison', a philosophical question now presents itself: are most bison not really bison? Should managers of conservation herds cull their hybrids and replace them with uncontaminated animals that have no 'introgression' of cattle genes? And does it matter if some 'real' bison genes are lost in the process, as some of these original genes can now be found only in individuals that are part cattle?

For Derr, these questions aren't hard. Although behaviour and morphology are important, Derr contends that a species is its genome. "If you don't have the genome, nothing else you do makes a damn difference," he says. "What you are preserving isn't the species; it is something the hell else — a shadow."

There is no obligation to save every version of every gene, but wildlife managers should keep pure herds from mixing with those that have cattle genes, he says. It is a simple formulation that has caught on with a number of conservationists.

"Jim has been the lead bull in bringing about a rigorous consideration of the role of genetics in bison conservation," says Kent Redford, head of the Wildlife Conservation Society Institute

in New York and director of the American Bison Society. "There have been those who have not agreed with Jim but the general movement has been towards the importance of his point of view."

— Rurik List

**"There are more important things than genes."**

And he "would not have had the success he has had without his personality", Redford adds. "He is charming, slightly intimidating, tireless and single-minded. I think it helps that he is most definitely a Texan."

But not everyone is convinced. "There are more important things than genes," says Rurik List, an ecologist at the National Autonomous University of Mexico, who works with a herd that spans the US–Mexico border. These bison have some cattle genes, but they also have



**When bison populations crashed, the New York Zoological Society and William Hornaday (right) helped save the species.**

U.S. FISH & WILDLIFE SERVICE



SMITHSONIAN INSTITUTION ARCHIVES

institutional memory. If List were to remove them and replace them with pure animals, would the bison still be able to find the water holes that the current herd knows so well? "They have been behaving like bison for 80 years," says List. "They have been fulfilling an ecological role."

Greg Wilson of the Canadian Wildlife Service in Ontario agrees. "I am less adamant about the pure bison thing," he says. "Whether you have 0% introgression or 2% introgression, it seems pretty arbitrary to me."

Until now, Derr has used about 14 loci to look for cattle genes in bison herds — that is, 14 places in which a certain pattern of base pairs will signify a cattle ancestor. The loci were narrowed down from 200 candidates in areas on the genome that have had a relatively brisk level of mutation in mammals. Derr found short spans, each of about 100–200 base pairs, that gave a clear bison or cattle signal and that were far enough away from each other to constitute independent observations.

However, the function of genes that overlap with or are near to these loci is unknown. So far, the only established difference between bison with and without cattle genes is their weight, a common indicator of the overall condition of an animal. According to Derr's as yet unpublished work, 5-year old male bison with cattle mitochondrial DNA were on average 4.5 kilograms lighter than those without.

But with just 14 loci, Derr hasn't been able to say anything about individual animals. A bison with 1% cattle genes might have them in places besides the 14 loci that Derr tests. Statistically, he needs a sizeable sample of animals — perhaps many dozens — to say for sure whether the herd has cattle genes in it.

Now that the domestic cattle genome has been sequenced, Derr is working on a 'SNP chip', a ready-made test to identify thousands of individual nucleotides that will be able to tell what proportion of the bison genome is made up of cattle genes. All he needs to do to find these nucleotides is to compare the two genomes and see where the differences are. He says he needs

US\$150,000 and a year and a half to make it work, and that each test will cost less than \$75.

Derr is a persuasive man. He has won over some government officials, but whether his efforts will translate into policy changes is uncertain, especially as many of the decision-makers are likely to be replaced under the administration of President Barack Obama. Kaush Arha, the former deputy assistant secretary for Fish and Wildlife and Parks at the Department of the Interior put together a bison working group that met with Derr and other scientists in October. Arha predicts that policy will eventually enshrine Derr's thinking. Moving animals from a hybrid herd to a clean herd, for example, might become forbidden.

Peter Dratch, a specialist in endangered species at the National Park Service, agrees. "We are now in an era where we need to look at the genetics of species," he says. For the bison, he says, "introductions to a herd should take place only when they do not increase overall levels of cattle ancestry". At the same time, small, pure herds that don't have enough animals to have healthy gene pools should be moved around so as to keep the genes flowing. "We will probably have to artificially create immigration."

### Defining features

Dratch even floats the idea of differentiating between 'wild' bison and those that have been commercially bred, which might be better thought of as livestock. "A wild bison," he says, "is a member of a herd of significant size with low levels of cattle or subspecies introgression that is subject to forces of natural selection — including predation, and competition for forage and breeding opportunities."

Such a definition might have implications for the Endangered Species Act, which is notoriously unclear about hybrids. If bison with more than a certain proportion of cattle genes don't count as bison under the act, and if, as Derr suspects, only about 8,000 animals have

**"Species conservation is more than skin deep."** — James Derr

no detectable cattle genes, then some bison might even qualify as endangered.

The American Prairie Foundation took Derr's ideas seriously when it went shopping for bison to place in its reserve in central Montana. So did the Nature Conservancy when it came to stocking a preserve in Iowa. Both organizations chose bison from Wind Cave National Park in South Dakota.

The Wind Cave herd is one of only two federally owned herds that Derr is confident contain no cattle genes. To a visitor, the pure bison at Wind Cave seem remarkably tame. They will approach a stopped car to lick road salt off of its undercarriage, and seem unperturbed by camera-wielding humans.

It may be that when Derr's SNP chip is ready some cattle genes will be found in the animals at Wind Cave. After all, they are separated only by a chain-link fence from bison known to have cattle genes. But if introgression levels are low, individual animals with cattle genes could be plucked out.

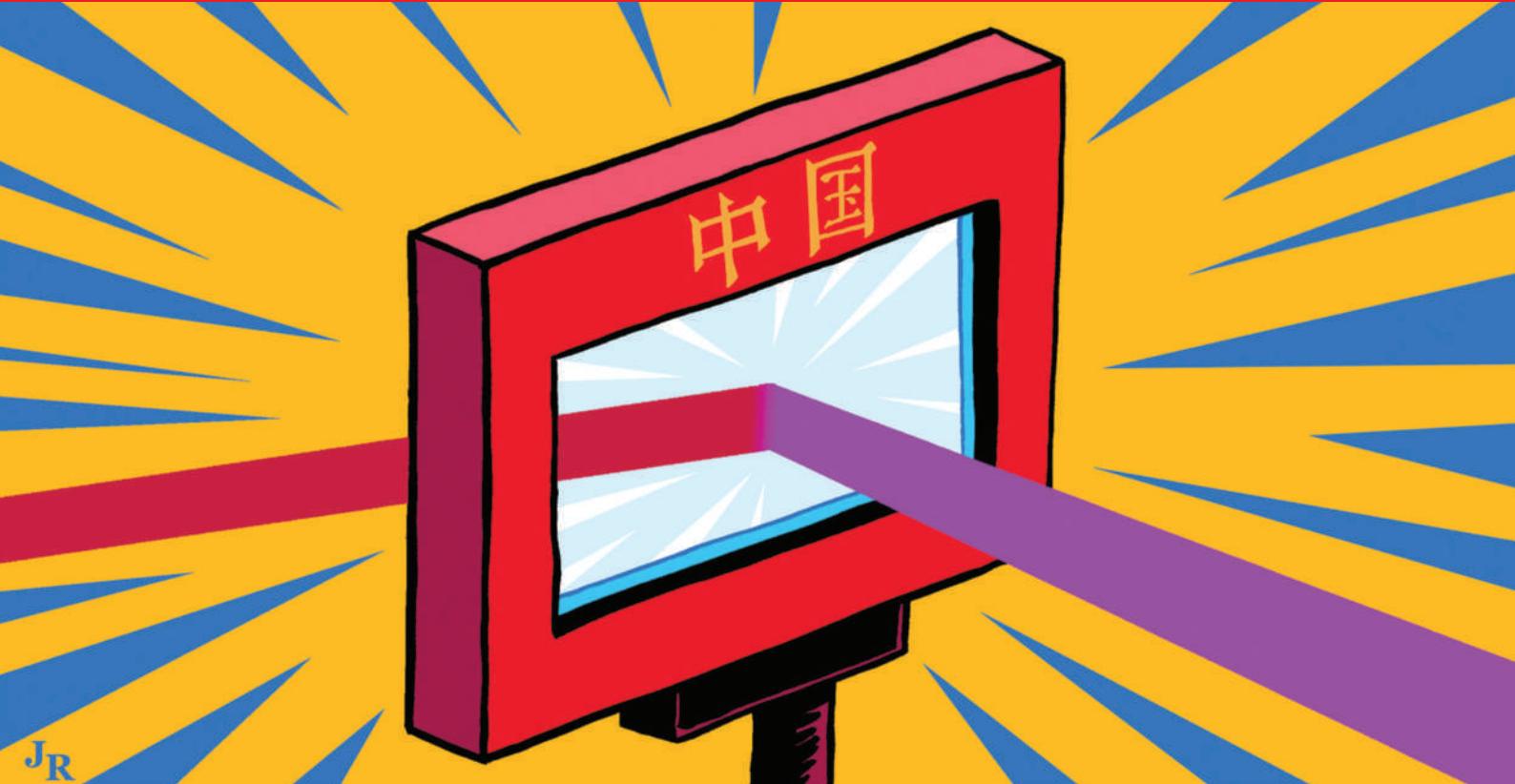
Derr hopes that conservationists will use the sequences from domestic animals to learn about the genomes of wild species that have chequered genetic pasts. Some hybrids have arisen for economic or accidental reasons.

For example, cattle have bred with the massive wild gaur (*Bos gaurus*) and Southeast Asian banteng (*Bos javanicus*). Others were created for conservation purposes: birds such as the peregrine falcon, for instance, have been bred across subpopulations for reintroduction into the wild. With a toolbox of SNP chips, wildlife managers may be able to tease apart mixed-up lineages and preserve pure genomes. Or such tests might show that it would be detrimental to try to weed out domestic genes from wild populations.

But the genomic approach goes beyond just working with hybrids. Looking at species as the sum of their genetic variation could help near-extinct populations. It could help managers to decide whether to bring two isolated groups of the same species together to improve genetic variation or to keep them separate. Suddenly, the world's fauna stops looking like a bunch of animals, and seems more like a group of genomes riding around in animal-shaped containers.

Derr is working on a number of genomes, but has a special place in his heart for bison. They symbolize everything about the West, he says. "Bison are incredible survivors." ■

**Emma Marris writes for *Nature* from Columbia, Missouri.**



JR

# China's crystal cache

A Chinese laboratory is the only source of a valuable crystal.  
**David Cyranoski** investigates why it won't share its supplies.

**O**ne of Daniel Dessau's prized possessions is a small crystal of potassium beryllium fluoroborate (KBBF). Dessau, a solid-state physicist at the University of Colorado at Boulder, uses the crystal to convert the light of a US\$100,000 laser into a deep ultraviolet, a good wavelength for studying the surface of superconductors. But because the laser light gradually degrades the crystal, Dessau has to save it for special projects. "It is a beautiful crystal," he says. "It would really move the field forward — if people could get it."

But Dessau can't get any more of it. Nor can Peter Johnson, a condensed-matter physicist at Brookhaven National Laboratory in Upton, New York, who was once promised it by Chuangtian Chen, the Chinese physicist who runs the only laboratory that knows how to make the crystals. And nor can any of a host of other solid-state physicists outside China. "There has been a limited release," says Johnson. "I don't know the politics behind it."

In fact, the politics is simple. The Chinese government is squeezing the crystal for every bit of academic and, eventually, commercial potential it can yield. In October 2008, the finance ministry sidestepped traditional scientific funding channels and started throwing 180 million renminbi (US\$26 million) at a three-year national project to find better ways to produce and use KBBF. China has selected a

handful of groups to work with Chen's crystal, including teams studying the newest type of superconductor, called pnictides.

China's monopoly of this crystal is no fluke. At a time when materials scientists and solid-state physicists elsewhere are seeing a lack of investment, their counterparts in China are surging ahead in a wide range of materials

research for much the same reasons as they did with KBBF. The nation has accumulated a great depth of crystal-growing know-how over the past three decades; it has steadfast government support; and its scientists are willing to subsume

themselves in a large team effort and take on the often thankless, sometimes dangerous and always tedious trial-and-error task of synthesizing new materials. "Many great discoveries in this field come from putting things together and getting the temperature and timing just right," says Christos Panagopoulos, a materials researcher at Nanyang Technological University in Singapore. The discovery process "doesn't require genius", he says.

KBBF's ability to shorten the wavelength, and thereby boost the frequency, of laser light is an example of 'nonlinear' optics, a field that first blossomed in the 1960s as lasers became more widespread in laboratories. Under ordinary

circumstances, light passing through water, glass or any other material will perturb the atoms only slightly, so that they vibrate in sync with the light wave. As a result, light can be reflected, refracted, scattered and absorbed ad infinitum without its frequency being affected. Nonlinear effects are evident only when the light is so intense that the vibrations it causes

compete with the binding forces on the atoms. When highly perturbed, as in the case of high-intensity lasers, the atoms can absorb the energy of the incoming light and re-emit the light with a frequency that is double, triple or even some

higher multiple of the original. A variety of materials have been discovered that can boost laser light to frequencies that the lasers alone cannot produce, and each has a set of signature frequencies that it can achieve.

China might easily have fallen behind in this field, as it did in so many others. Just as nonlinear optics started coming into its own, China was caught up in the Cultural Revolution, a particularly dark period starting in the mid-1960s when many academics were criticized as being elitist or impractical and sent to do farm work for 're-education'.

But Chen, now a sprightly 71-year-old at the Technical Institute of Physics and Chemistry

**"You need a lot of equipment and you need to move slowly."**  
 — Chuangtian Chen

in Beijing, was lucky. "The government always considered crystals important for industry," he says. So by agreeing to the government's request to switch from its theoretical studies to growing crystals, Chen's lab was able to continue doing materials science throughout this period.

### Crystal blueprint

In the 1970s, Chen developed a formula that set out the conditions needed for a material to generate nonlinear effects. In 1984, the formula led Chen and his team at the Fujian Institute of Research on the Structure of Matter to investigate barium borate ( $\text{BaB}_2\text{O}_4$ ), which proved to be the first material able to generate an ultraviolet wavelength of close to 200 nanometres. The crystals are now widely used in femtosecond lasers, which use ultrashort bursts of infrared light to slice through materials with extreme precision, making them ideal for some types of surgery.

This was before China had any laws covering intellectual property, so Chen received no royalties, but his salary jumped from 87 renminbi a month to 147 renminbi a month. Chen thought he had hit the big time.

Another seminal discovery came in 1987, when Chen's group demonstrated nonlinear optics in lithium triborate ( $\text{LiB}_3\text{O}_5$ ). Engineers now use the compound to create high-powered green and near-ultraviolet lasers. As the crystals are extremely resistant to damage, they are particularly useful in applications such as welding and semiconductor manufacturing. Some 80–90% of all solid-state lasers and laser systems now use lithium triborate crystals for frequency

conversion. Castech, a corporate spin-off set up by the Fujian Institute of Research on the Structure of Matter to manufacture and develop crystals for use in lasers, has been making several million US dollars a year from the compound.

But Chen and his PhD student Rukang Li were not done shortening wavelengths. Starting in 1988 on what would be a long journey, Li examined the chemistry of "hundreds, if not thousands of compounds" using Chen's formula. "KBBF looked good," remembers Li. Eight years later, they proved that the crystal could produce a laser of an unprecedentedly short, 184.7-nanometre wavelength<sup>1</sup>.

However, the crystal was still not ready for practical use. KBBF grows thin and plate-like, making it difficult to cut at the angles needed in lasers. It took another seven years before a collaboration set up in the late 1990s between Chen's group and Shuntaro Watanabe at the University of Tokyo succeeded in getting KBBF into a laser system<sup>2</sup>.

That is when scientists really started to get excited. Exploiting the fact that the ultraviolet light from a KBBF-equipped laser has an extremely narrow frequency range, allowing it to measure the energy level of electrons in solids down to a resolution of just 360 micro-electronvolts, Shik Shin from the University of Tokyo and his colleagues were able to show that the fine spacing of energy levels in certain superconductors depends on the direction in which the electrons travel through the lattice<sup>3</sup>. "It was the first time I could really say I was a pioneer, that I was seeing something that nobody else had seen," says Shin.

That discovery, the first to be made with a KBBF laser, opened up investigations of many different kinds of superconductors that had been impossible before because there was no laser available that had sufficient energy resolution. Since then, Shin and Chen have co-authored more than 20 papers.

Xingjiang Zhou, of the Institute of Physics in Beijing, was also using KBBF to examine superconductors. He discovered an entirely new type of electron pairing<sup>4</sup> — the fragile coupling that allows the electrons to move through the lattice without resistance. One leading condensed-matter physicist, seeking a collaboration, told Zhou in an e-mail, "These are the highest quality data that I have ever seen."

### Clamp down

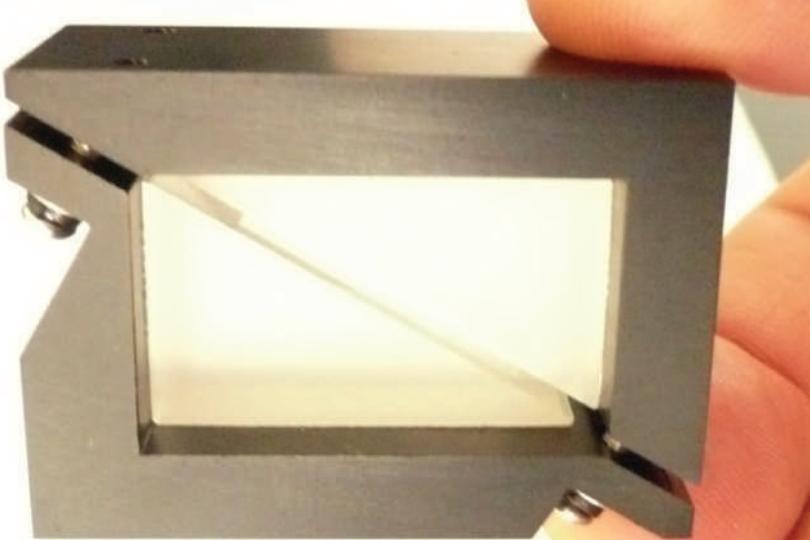
By 2008, thanks to these and other stories, requests for KBBF crystals were rolling in to Chen's institute. And that, in turn, caught the attention of the institute's parent organization, the Chinese Academy of Sciences, which told Chen not to distribute the crystal any further without its permission.

Chen is planning to use his share of the government's 180 million renminbi to install more large ovens in which to grow crystals. This will allow his institute to ramp up from the 15 KBBF-crystal devices it made in 2008 to 50 in 2009, and then 100 in 2010. His team will also be looking for ways to produce better KBBF crystals. Thicker crystals allow for a more powerful laser and make possible hugely profitable applications, such as replacing the bulky exciplex lasers, another type of ultraviolet laser used in surgery and in semiconductor lithography. Chen, who is discussing the possibility of technology transfer with two companies in Beijing, hopes to find some commercial applications within three years.

Seven other projects will be aiming to create advanced versions of photoemission spectrometers, Raman scattering spectrometers and scanning tunnelling microscopy. Zhou will receive 20 million renminbi to head two of the projects. In one, he plans to make a tunable KBBF laser that can analyse a wider variety of materials — at present, these lasers need to be set at one frequency. In the other project, he will develop a KBBF-based laser to look at the spin of electrons in superconductors. Until now, research has focused on momentum and energy.

Chen is reluctant to talk about the terms of the government's restriction. He says he would like to share the crystals with people in other countries, but first has to meet demand from Chinese KBBF projects. "The government gives me so much money," he says.

Crystal growers in other countries are unlikely to be able to fill the gap, mainly because of the



**KBBF sandwich:** it took Chen's team nearly 15 years to grow KBBF and put it into a usable form.

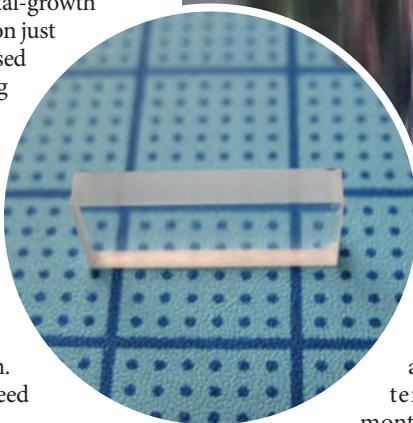
time, people, infrastructure and know-how needed to create a good KBBF crystal. Chen's laboratory has 70 people, including technicians and some 30 students. The group has learned how to make KBBF using the same factory-like process they use to develop all their other nonlinear crystals — calculating what compounds might work; synthesizing the material; growing, cutting, polishing and coating the crystals; then testing them for their vulnerability to damage by intense laser radiation and other properties. Chen's group has put in huge resources for infrastructure such as crystal-growth ovens. It has spent US\$7 million just on platinum for the crucibles used to grind up the crystals during the production process. Even now, with the KBBF process honed, it takes 3 months and about US\$20,000 to produce just one KBBF device. Bruce Chai, president of Crystal Photonics in Sanford, Florida, says that anyone trying to duplicate the work would easily burn through US\$10 million. "And they wouldn't be guaranteed success," he adds.

Perhaps the greatest obstacle to making a KBBF crystal is the beryllium it contains, which can cause pneumonia-like symptoms and cancer if inhaled. China has strict policies regarding its use, says Chen, but researchers are allowed to work with the element if they have the right infrastructure. "You need a lot of equipment and you need to move slowly," he says. In other countries, the restrictions are more severe. In the United States, for example, lawsuits over beryllium poisoning led the Department of Energy to hunt down and remove even trace amounts from national laboratories.

"The demand for KBBF is spurring researchers to look for other fluoroborates that present fewer challenges," says Vincent Fratello, vice-president of research and development at Integrated Photonics in Hillsborough, New Jersey. But until that search bears fruit, the benefits of Chen's dogged pursuit will stay in China.

KBBF is a special case, but it illustrates China's growing strength in materials science. Fratello, himself a solid-state physicist, says that he has been impressed by the "sheer volume of work that comes out of China".

Fudan University's Dongyuan Zhao has more citations than anyone in the field of mesoporous materials (and the second highest number of citations per article). Last year, after Japanese researchers discovered high-temperature superconductivity in iron-based arsenic oxides, several Chinese groups jumped



C. CHEN

C. CHEN

Government investment has helped Chuangtian Chen to develop KBBF crystals (inset) for use in lasers.

in to investigate, achieving even higher temperatures within months. Zhou says that he intends to use the KBBF laser to take arsenic oxide studies even further.

### Material leads

Some researchers see China as following the lead of Japan, where significant investment in materials sciences in the 1990s has been paying off in discoveries such as the superconductor magnesium diboride<sup>5</sup>. Panagopoulos, who last year moved from the University of Cambridge, UK, to Singapore to continue his research on functional materials, says he tried and failed to get Cambridge to bulk up its materials-synthesis capacities in the late 1990s. "We wait for Japan, and now China, to make it," he says — a sure recipe for mediocrity. "The person who has the greatest knowledge of everything about a material is the one who is distributing it to everyone," says Panagopoulos.

There is also concern in the United States. "It is a well-known fact that the support for basic research on crystal growth is literally gone in the United States," says Chai. This spring, the US National Academy of Sciences will release a report on the health of, and future opportunities for, new materials and crystal-growth research.

Even in trendier, related fields such as

solid-state physics, China is catching up with the traditional leaders. High-temperature superconductor specialist Hong Ding had several attractive offers last year. But neither Boston University in Massachusetts, where he had been for a decade, nor any other institution could match the deal he was offered at the Institute of Physics in Beijing. "It is a matter of time before the United States becomes alarmed by this rapid reverse of the brain drain," says Ding. Dessau, who tried and failed to recruit Ding, says that "10 years ago it would have been unheard of [for a Chinese person to turn down a position in the United States]. But I wouldn't be surprised if the trend continues."

In fact, many credit the hard-working, selfless laboratory worker for some of China's success, and Japan's before it, in rapidly increasing their capabilities to synthesize materials. Fratello says that Chinese groups

"tend to excel at studies that require patience and looking at a lot of different systems". But in materials sciences, labour-intensive work can often be the most productive. "That is, after all, how you discover new materials," Fratello says.

**David Cyranoski is Nature's Asia-Pacific correspondent.**

- Chen, C., Xu, Z., Deng, D., Zhang, J. & Wong, G. K. L. *Appl. Phys. Lett.* **68**, 2930 (1996).
- Togashi, T. et al. *Opt. Lett.* **28**, 254–256 (2003).
- Kiss, T. et al. *Phys. Rev. Lett.* **94**, 057001 (2005).
- Zhang, W. et al. *Phys. Rev. Lett.* **100**, 107002 (2008).
- Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. *Nature* **410**, 63–64 (2001).

# CORRESPONDENCE

## Scientists on Turkey's banknotes should inspire young minds

SIR — Countries that have adopted the euro as currency no longer commemorate their national identity with famous cultural figures on their banknotes: a sad loss. Gone, for example, are great scientists such as Carl Friedrich Gauss, who was portrayed on Germany's 10-mark note, Pierre and Marie Curie from France's 500-franc bill, and Alessandro Volta, with the demise of Italy's 10,000-lira note. For different reasons, Yugoslavia's ten-billion-dinar note honouring Nikola Tesla has also disappeared.

Happily, 2009 is witnessing a revival of 'scientific' banknotes — and not just in the United Kingdom, where the Bank of England's £10 note reminds us of the bicentenary of Charles Darwin's birth. The Republic of Turkey's latest banknotes depict Turkey's founder Mustafa Kemal Atatürk as before, but on the back they show portraits of notable intellectuals. For example, the Turkish 5-lira note commemorates Aydin Sayili (1913–93), a science historian who won a state-supported scholarship under George Sarton at Harvard University as a result of a chance meeting with Atatürk. The same banknote carries schematics of the Z-form of double-stranded DNA, rich in G-C base pairs, and a beryllium atom with orbiting electrons, both of which are echoed on a hologram foil strip on the right of the face sheet. The carmine 10-lira note portrays the Turkish mathematician Cahit Arf (1910–97) — known for his invariant, which is applied in knot theory and surgery theory — with geometric figures that are also reproduced on the hologram foil strip.

This departure reflects the significance paid to science and educational development in Turkey. The new notes delight scientists but they should also inspire young minds — in contrast

to the depersonalized euros of a unified Europe, which just indicate denominations of currency.

Turkey has more than 100 universities, both state and private, most of which were set up during the past 15–20 years. These employ some 10,000 professors, many of whom have worked abroad. Against this thriving background, it is anticipated that Turkey's new banknotes will retain their value for a long time to come.

**Renad Zhdanov** Department of Genetics and Bioengineering, Yeditepe University, Istanbul 34755, Republic of Turkey  
e-mail: zhdanov@yeditepe.edu.tr

## Protecting the Hawaii akepa population

SIR — Your News story 'Feathers fly over Hawaiian bird' (*Nature* **456**, 682–683; 2008) raises important questions about the population status of the endangered Hawaii akepa (*Loxops coccineus*) in the Hakalau Forest National Wildlife Refuge in Hawaii. It focuses on the research of Leonard Freed and his team on Hawaii's endemic birds, the implications of which have been seriously considered by the US Fish and Wildlife Service (USFWS).

USFWS staff, including refuge biologist Jack Jeffrey, took several actions to ensure that Freed's research did not fall victim to differences of opinion or personality. The Hawaii Forest Bird Recovery Team was asked to review Freed's work and its management implications on multiple occasions. Also, the USFWS initiated a review of the population status of the akepa by independent scientists and hosted a workshop in October 2008, in which Freed participated, to identify research and management priorities for the refuge.

The consensus of these reviews was that the akepa population

is showing no signs of imminent collapse and faces greater threats than competition from the Japanese white-eye (*Zosterops japonicus*). The majority of the 21 scientists participating in the workshop disagreed with Freed's conclusions for several reasons.

First, they had a contrasting assessment of the population status of the akepa, based on annual abundance estimates taken over the past 21 years. These data did not suggest a population crash or a negative association with the Japanese white-eye.

Second, the scope of inference from annual surveys differed from that of Freed's work. Annual surveys were based on 300 stations established using a probability sample of the entire 13,400-hectare refuge, allowing inferences to be made for the whole refuge. Freed's conclusions were based on observations from a few small study sites within the refuge, undermining the extrapolation of his results to a refuge-wide phenomenon.

Given current budgetary challenges, management actions such as habitat restoration and the removal of hooved animals that damage habitat are the top management priorities.

**J. Michael Scott** US Geological Survey, Department of Fish and Wildlife, University of Idaho, Moscow, Idaho 83844-1141, USA  
e-mail: mscott@uidaho.edu  
**Jon S. Horne, Edward O. Garton** Department of Fish and Wildlife, University of Idaho, Moscow, Idaho 83844-1141, USA

## Evolutionary gems of the plant world shine just as brightly

SIR — I applaud *Nature's* initiative in bringing together '15 evolutionary gems' (*Nature* **457**, 8; 2009; [www.nature.com/evolutiongems](http://www.nature.com/evolutiongems)) that provide empirical evidence for the process of evolution by natural selection. But I was struck by the

conspicuous absence of plants (not to mention invertebrates and microorganisms) from the list, which is intended for the enlightenment of non-biologists. This risks sending a wrong message, such as that there is insufficient evidence for plant evolution (or, worse, that plants are not important), when in fact many of the benchmark contributions to the understanding of plant evolution have been published in *Nature*.

Sadly, 'plant blindness' — a lack of awareness of and interest in plants in biology education and among the general population — is well documented (see, for example, E. E. Schussler and L. A. Olzak *J. Biol. Educ.* **42**, 112–118; 2008). But plants are key components in ensuring continuity of life on Earth. Their evolution intertwines at many points with that of animals.

Darwin's ideas on evolution were in part based on, and in turn influenced, his study of plants. His "abominable mystery", the perceived rapid diversification of flowering plants, still remains an important question in evolutionary biology (*Am. J. Bot.* **96**, 1–381; 2009).

*Nature* is well equipped to join the fight against 'plant blindness'. How about '15 more evolutionary gems' to ramp up awareness of the evidence for evolution in the plant (and invertebrate, fungal and microbial, for that matter) branches of the tree of life?

**Alexandru M. F. Tomescu** Department of Biological Sciences, Humboldt State University, Arcata, California 95521, USA  
e-mail: mihai@humboldt.edu

Readers are welcome to comment at <http://tinyurl.com/d8ctnv>

Contributions may be submitted to [correspondence@nature.com](mailto:correspondence@nature.com). Please see the Guide to Authors at <http://tinyurl.com/373jsv>. They should be signed by no more than three authors and preferably by one. Published contributions are edited.

## COMMENTARY

ILLUSTRATIONS BY DANIEL MACKIE



RECESSION WATCH

# HOW TO SURVIVE THE RECESSION

The global economic downturn brings both predicament and promise.

How will science fare and what role should scientists play on the long road back to recovery and growth?

Ten of the world's leading thinkers and practitioners provide analysis, experience and advice.

**COMMENTARY**

958 **Boost the developing world**  
Jeffrey Sachs

958 **Learn to convince politicians**  
Ian Taylor

959 **Work for the greater good**  
Eric Rauchway

960 **No time for nationalism**  
Atsushi Sunami &  
Kiyoshi Kurokawa

961 **Cut costs and sell what you can**  
John Browning

962 **Cooperation must rule**  
Noreena Hertz

963 **End the obsession with interest**  
John Geanakoplos

**BOOKS & ARTS**

964 **Old lessons for a new economics**  
Bill Emmott

965 **Investing in the environment**  
Gail Whiteman



For podcast and more online  
extras see <http://www.nature.com/recessionwatch>

# Boost the developing world

**Directing finance into sustainable infrastructure in the poorest countries helps the whole world, says Jeffrey Sachs.**

The economic crisis that originated in the consumption boom and bust of the United States and Europe has quickly spread, and is now hitting developing countries hard. This requires urgent global attention: inhabitants of poor countries lack the personal and societal buffers that enable those in richer countries to weather the economic storm that they have created. But the crisis also provides global opportunity. A financial package from the G20 — a forum of cooperation among 20 large economies — to support sustainable energy, land and water use in the poorest countries could provide a ‘triple win’. It could lead to stimulus for richer countries, development for poorer countries and environmental sustainability for all.

Sub-Saharan Africa’s income growth in 2009 is projected at just 3.25%, down from 6.9% in 2007, and around 3 percentage points below the forecast of a few months ago. The Institute of International Finance based in Washington DC projects a collapse of private-sector financial flows to the world’s emerging market economies (mainly middle-income countries such as Morocco in North Africa), from US\$928.6 billion in 2007 to just \$165.3 billion in 2009. International banks are expected to reduce their loans to the middle-income countries by around \$60 billion, compared with a net loan increase of \$410 billion in 2007. Money sent from expatriates to their families in poor countries is plummeting, and millions of workers are returning home after losing jobs, work permits and visas. Long-promised aid is already being cut by some of the largest donor countries. Italy, the host of the Group of Eight (G8) Summit this year, has cut its 2009 aid budget by around half.

These results, compounded by the effects of changing climate, could well prove calamitous. The world’s hungry population has increased by 100 million or more compared with 2003–5. Famine stalks large parts of Africa as a result of pervasive climate shocks, growing financial crisis and political instability. The rich world has



looked on, made ample aid pledges in earnest speeches and summit texts, and so far delivered almost nothing of what has been promised.

The tragedy of inaction is made worse by two basic truths. First, the amounts promised but not delivered for Africa by the rich world — about \$60 billion per year up to 2010, compared with actual aid flows stuck at around \$30 billion — are an insignificant fraction of the trillions of dollars of financial bailouts and stimulus packages adopted by the rich countries in the past four months. Merrill Lynch’s Christmas bonus of \$4 billion, using bailout funds, was only slightly less than the \$5 billion of US aid to Africa.

Second, the United States, Europe and Japan are missing out on a huge opportunity for themselves: should they actually follow through on their unmet promises of grants and loans to help build Africa’s roads, power-generation facilities, ports, water and sanitation systems, fibre-optic networks and functioning agriculture systems, they would provide an enormous stimulus for the faltering factories of the donor countries. Ironically, there is a glut of solar-panel production throughout the rich world, because domestic demand has dropped, while at the same time an absolutely urgent need for solar power in the tens of thousands of African villages not connected to a power grid. That solar power would enable Africa to pump clean water, irrigate crops, run grain-grinding mills, and power clinics and hospitals.

The G20 countries, meeting in London in early April, would be wise to devote at least \$25 billion in urgent additional funding for African sustainable investments in 2009, and another \$25 billion for low-income countries in other regions. The numbers can hardly shock in a meeting that will spend all its time discussing trillions of dollars for our own economies, but

whether the G20 will spare attention and money for the world’s poorest is anyone’s guess.

The rich countries frequently tell the poor that “it’s not the lack of aid but your own governance” that slows the pace of development, even as many poor countries have greatly strengthened their governance and are held back by insufficient aid. Such words ring even more hollow today, as the poor reel under the weight of rich-country scandals, regulatory breakdowns, financial crises and the impacts of climate change emanating from the rich world. All of the world will be losers if the needs of poor countries remain unaddressed, causing rising violence, disease, population displacement and shrinking markets. ■

**Jeffrey Sachs** is director of the Earth Institute at Columbia University, New York, USA and author of *Common Wealth: Economics for a Crowded Planet* (2008).  
e-mail: sachs@columbia.edu

## Learn to convince politicians

**Scientists must be prepared to explain why research budgets need protecting when times are tough, says Ian Taylor.**

The severity of the current economic downturn is much greater than when Britain last experienced a recession in the early 1990s. Yet there may be lessons to learn from that downturn. Now more than ever, government has a role in focusing strong, directed efforts to boost the translation of scientific ideas into useful technologies, and to reinforce the base of science skills that drives this innovation.

The United Kingdom cannot rely on the formerly dynamic financial services sector as the engine of growth, so a rebalancing of economic activity is vital for recovery. The ability to capture ideas and discoveries that flow from research will be the test of whether we can recover growth and prosperity, and will influence the degree of public support for investment in science.

I was UK science and technology minister in 1994–97, as we were emerging from the last recession. The 1996 budget was a vigorous attempt to rein in public expenditure,

which had been rightly boosted during the early 1990s. Everyone was asking who had performed well during the tough times, and which budgets could now be cut. I held off efforts to raid the science budget, but scientists themselves were not the most helpful in persuading my fellow ministers. Scientists thought it was self-evident that they should remain immune to cuts, but too often failed to articulate how they could help the community through tough times. I accused one organization of "whingeing" on public radio, as that is how it sounded.

There have been improvements since then. The Royal Society, for example, has become more realistic and more focused on the practical. I now sit on the Advisory Council for the Campaign for Science and Engineering — the organization I accused of moaning. But we need to go further.

Applied science should be raised to the status of basic science. From my visits to universities, I know how many scientists are frustrated that they are not given financial incentives to take projects beyond publication of an outcome. Yet this is exactly what is needed to help solve today's most pressing problems.

In the early 1990s, the Conservative government launched a lot of worthy schemes, including some to stimulate small spin-off companies. But we were too scattergun in our approach, and missed the point slightly by focusing on ideas coming out of universities as if innovation always followed a linear path.

We have learned from that. The Conserva-

tive Party's Science, Technology, Engineering & Mathematics task force, which I chair, proposed the creation of a powerful Innovative Projects Agency (IPA) alongside the Research Councils to refocus spending on innovation into areas of national interest, and to better bridge the gap between the birth of an idea and the development of a commercially attractive technology. The IPA would recognize that good work comes not just from university labs but from industry too, and would encourage collaboration and creative engineering. It would recognize that innovation entails uncertainty, as opposed to quantifiable risk. This would mean a radical shift in the government approach to science funding and innovation, with more stress on fulfilling demand rather than to funding supply.

What are the key problems of national interest that need addressing? Revolutionizing transport technologies, meeting climate-change targets and securing diversity of energy supply, to name a few. Both nuclear and clean-coal technologies require more directed research, and more courses to teach these skills. The United Kingdom does not at present have the workforce needed to enact these technologies.

The country is generally exposed to risk by not producing enough domestic talent. A recession creates opportunities in this regard. Surveys show that science students are more employable: the Leitch Review of Skills predicts that the demand for science and technology professionals and associated staff will rise by 18% and 30%, respectively, between 2004 and 2014, compared with 4% for all other occupations. Now is a great time to increase recruitment into science study.

It is also a good time to reverse the outflow of graduates from careers in science. In 2002, 6% of physics graduates entered the world of finance; by 2007 that was up to 19%. The downturn in the financial services sector should make it easier to keep science students in science.

When we inevitably come out of this financial crisis, the science and engineering communities will need to be able to highlight their practical successes during this period, not least to protect their funding as the government claws back on public spending to reduce debt. This means unleashing a real enthusiasm in the science community to facilitate more mission-driven science, and real support from the government to allow that to happen. ■

**Ian Taylor**, a member of the UK parliament, chairs the Conservative Party's task force on science, technology, engineering and mathematics, and is a member of the Advisory Council of the Campaign for Science and Engineering. e-mail: taylori@parliament.uk



## Work for the greater good

During the Great Depression, scientists proved to America why researchers are key to nation-building, says **Eric Rauchway**.

When hard times hit, research budgets shrink. But if the US response to the Great Depression of the 1930s provides a guide to the current crisis, scientists may find a way to benefit from reform policies. After an awkward start, the New Deal programme established a new era in government-science relations, with expanded funding for basic research and newly coordinated research efforts to support goals such as resource conservation and public health. Eager for solutions and amenable to what president Franklin Roosevelt called "bold, persistent experimentation", policy-makers heeded scientists with practical ideas of how to use research to solve social problems.

In 1933, Franklin Roosevelt created a Science Advisory Board, the members of which hoped to run a national research programme. Like other central-planning agencies of the early New Deal, it failed to attract sufficient ongoing political support and was moribund by 1935, when Washington reaffirmed



its commitment to receiving advice from the National Academy of Sciences. But in the wake of this vague failure ran numerous successes.

Chief among these was the Tennessee Valley Authority (TVA), charged to develop the chronically impoverished and acutely over-farmed Tennessee River watershed. Best known for its hydroelectric dams, the TVA also employed scientists to improve the well-being of the Old South. For Harcourt Morgan, the agricultural scientist who became TVA chairman, a given innovation had not only to yield near-term economic efficiency, but also to promote sustainable use of resources.

TVA scientists ran experimental farm stations to cultivate crops chosen to prevent soil erosion. They reforested thousands of acres and resettled the region with deer. They reported on water quality, to shame polluters, and controlled water-borne diseases. The TVA built on the federal government's long-standing research programmes in the US Department of Agriculture, but it lent a newly coherent vision to the role of agricultural research in preserving American resources, and implemented innovative techniques accordingly.

In 1935, the new Works Progress Administration (WPA) brought these regional policies to the nation. Primarily meant to employ the jobless, the WPA funded laboratories at public universities and all kinds of basic research, as a quick look through scientific journals of the later 1930s reveals: studies of foraminifera, *Drosophila* and coliform bacteria; archaeology, geology and biology all enjoyed WPA assistance, which ranged from funding to providing statistical analysis to simply supplying willing workers. The WPA also promoted conservation, where again scientists played a major part, as in the design and construction of the dunes to prevent the erosion of the North Carolina seashore at Cape Hatteras.

From 1934 the United States saw annual increases in government funding of research, so that by 1940 public money for science had increased three-fold over its 1934 level. The research, often aimed at increasing farm productivity, sometimes yielded unplanned dividends, as when mycologists at a New Deal-created research centre later developed techniques for the mass production of penicillin for antibiotic purposes.

None of these programmes was without cost, and with increased federal funding for science came now-familiar critiques. For example, TVA dams sometimes required the relocation of whole populations; its miracle fertilizers became pollutants; its provision of electricity to the Manhattan Project's Oak Ridge facility associated it with the ambivalence some Americans came to feel about the atomic programme.

Yet the overall successes were obvious and much copied. The TVA brought industry and modernity to the Old South, and the 1930s brought unparalleled technological progress to the nation, so that most industries saw improvements in productivity owing to technical innovations and their application. In the 1940s, the United States found leaders throughout the developing world eager to imitate New Deal programmes for the practical application of scientific progress.

If the current crisis is to yield similarly enviable models for scientifically driven economic advance, scientists with research experience and applicable ideas must speak out so the new administration can hear.

**Eric Rauchway** is professor of history at the University of California, Davis. He is the author of *The Great Depression and the New Deal: A Very Short Introduction* (2008).

e-mail: [earauchway@ucdavis.edu](mailto:earauchway@ucdavis.edu)

## No time for nationalism

**Basic research saw a boost in Japan's last recession. Better global links will help in the current one, say Atsushi Sunami and Kiyoshi Kurokawa.**

**D**uring the 1980s, Japan rose as a technology superpower. In the United States and Europe, policy-makers, journalists and academics talked about how to compete with Japan or even be like Japan. By the end of that decade, a bubble economy created by incredibly high land values and incredibly low interest rates was at its height in Japan, leading to massive borrowing and investment in stocks and securities — much the same as in the United States and Europe until recently.

Then the bubble burst. In 1990, the government imposed a restriction on loans for land purchase, hoping to control the rising real-estate prices. Instead, this led to a crash in the stock market and a crisis in the banking sector. This brought the Japanese economy to a com-



plete halt, which lasted throughout the 1990s: a period often called the 'lost decade'.

Yet it was in the middle of this lost decade that Japan created the foundation of its current science and technology policy.

The extraordinary sense of crisis brought together the best of Japan's scientists with the policy community — for the first time since the period of reconstruction after the Second World War. It was apparent that, in line with the country's rapid economic development, the bulk of Japan's research and development was being done by industry, rather than within universities and public-research institutes, and the government was convinced this needed to change.

Blue-skies research was seen as the way forward: a shining light pointing the way out of a dark tunnel. So in 1995, the Science and Technology Basic Law was introduced, the first of its kind in Japan. With the establishment of a new Council of Science and Technology Policy, headed by the prime minister, science had finally stepped centre-stage in Japanese policy-making.

The law served as the basis for Japan's first 'five year plan' for science and technology: from 1996 to 2000, public spending on basic and applied research was ¥17 trillion (around US\$160 million at the time), compared with ¥12.6 trillion for the previous five years. Spending targeted specifically to fundamental research increased from 1998, with good results.

Thanks to the example set by the first five-year plan, financial support for basic research was basically guaranteed. A second period, from 2001 to 2006, came with an investment of ¥24 trillion, and we are now in the middle of a third period with nearly the same level of spending. But that may not last. Now, the government in Japan, as everywhere else, is looking for places to downsize, and science is

no exception.

It is becoming ever more apparent that Japan, like other countries, faces global challenges that include the current economic crisis, climate change, poverty, health, and food and water shortages, to name just a few. Yet there is mounting frustration among the science-policy community that Japan's science system has taken the generous boost in funding without adopting the necessary reforms to meet such challenges head on. For Japan to tackle these problems requires not just money, but a more global view, and commitments that will allow Japan's research community to better compete, collaborate and thrive internationally.

Research grants should be made more accessible for international research projects, for example, so that foreign collaborators working in labs elsewhere find cross-border research with Japanese researchers easier. Students and younger researchers should be encouraged to go abroad for a period of study or for their early career, to improve their international networking when they later return to Japan. And Japanese universities and research institutions have to become more open and attractive to the outside world, in part by providing a significant fraction of courses in English.

More foreign students and researchers within Japan would both increase the overall standard and improve international links; currently, university enrolment of foreign students for undergraduate and postgraduate programmes is at about 3.5% in Japan, compared with 14.8% in the United Kingdom and 5.5% in the United States. Within universities, researchers should be given more academic freedom, incentives for competition, and help with technology transfer.

Some new policies have been adopted to address some of these issues: programmes to encourage collaboration and student exchange. Implementation should follow, but is not guaranteed — particularly if the science budget is cut. Yet such policies are necessary for Japan to be a leader in the global environment.

As in the 1990s, now is a time of crisis and an opportunity for positive change that leads to new, global common values. The last recession brought the country's first real science plan, which laid the groundwork for a boost in basic research funding. Now there is an opportunity for Japan to implement further reforms to become more globally competitive. ■

**Atsushi Sunami** is director of the science and technology programme and associate professor at the National Graduate Institute for Policy Studies in Tokyo, and **Kiyoshi Kurokawa** is a former special adviser to Japan's prime minister and a professor at the same institution.  
sunami-atsushi@grips.ac.jp

# Cut costs and sell what you can

Technology start-ups need to trim their sails in rough economic times, says **John Browning**.

The bell tolls in Silicon Valley. "RIP Good Times" was the epitaph that greeted elite Silicon Valley entrepreneurs — executives of companies invested in by Sequoia Capital of Menlo Park, California — at a meeting in October 2008. Sequoia is famous for having backed Apple, Cisco and Google, among others. In normal times, these executives would have been encouraged, as the valley's chosen ones, to focus on long-term strategic goals — market share, product development, international expansion. But Sequoia was telling them to forget all that. The recessionary motto is "spend every dollar as if it were your last".

Although this sounds like good advice for any time, the strict cash discipline imposed by recession actually represents a deep shift in management outlook. It raises tactics and short-term manoeuvring to equal importance with long-term strategy. It turns running a small company into something more akin to sailing a dinghy than driving a motor cruiser. Without financial engines to drive a company along whatever course the management should choose, success in a recession requires a constant search for the right combination of wind and current.

The dot-com bust of 2000 was littered with the wrecks of companies that failed to trim their sails in changing times. One of the most unnecessary losses was *The Industry Standard*, a weekly magazine based in San Francisco. *The Standard* was the first publication about doing business on the Internet, and it was also one of the fastest-growing publications ever launched. Revenue skyrocketed from US\$9 million in its first year, 1997, to \$200 million in 1999. The magazine's mistake was to eliminate its room for manoeuvre. It signed long-term leases and technology contracts. So when the downturn came, it couldn't cut back. Even though its reduced revenues should still have been sufficient for a profitable future, it foundered under inescapable high costs.

"The dot-com bust of 2000 was littered with unnecessary wrecks."



The first step in successfully steering a start-up through a downturn is to reduce costs early and dramatically. Most start-ups are cash-flow negative — that is, they spend more on product development and marketing than they take in revenues. Cutting costs gives them more time before they run out of cash. More time means a greater chance of outlasting the competition. Even Google, highly profitable and famous for allowing its engineers 20% of their time to work on new technologies of their choosing, began in autumn 2008 to cut 'dark matter' — speculative projects that aren't immediately catching on and no longer have much support.

But creating leeway to survive uncertainty is only half the struggle. Real success comes from the ability to turn uncertainty into opportunity. That has two parts.

One is to make more of every dollar of spending than the competition. The other is to grasp unexpected chances to gain extra revenue.

Many start-ups end up selling something other than what they started out to make. Flickr, a successful photo-sharing website, was originally created as part of a role-playing game. When usage of the photo-sharing facility took off beyond the scope of what it had been made for, the game was shelved and Flickr continued on its own. That kind of flexible thinking is especially valuable during a recession.

Even without a change of product, extra sales of early work can bring additional cash. Among biotechnology companies, it is a long-established strategy to sell diagnostics or contract research in order to support longer-term product development. Zappos, an online shoe

# Cooperation must rule

Navigating an unpredictable world will need different research disciplines to work together as equals, says **Noreena Hertz**.

store, has begun renting out its technologies for online sales, even to direct competitors such as Clarks. It reckons that any short-term loss of competitive advantage will be more than compensated by ploughing the extra resources back into developing long-term technological dominance.

On the spending side, one of the most fruitful places to search for increased efficiency is marketing and advertising. *Wired* magazine, launching on a shoestring in 1992 shortly after the worst of the 1990–91 recession, set a standard for advertising focus. It spent much of its advertising pittance on fluorescent-pink posters for the sides of the buses that trundled along Madison Avenue, the heart of New York's advertising industry. When *Wired*'s advertising salespeople came to call potential advertisers, they would be met with a flash of recognition rather than a blank stare.

Costs that can't be cut altogether can often be made more flexible, to go up or down with the company's fortunes. It is possible to rent rather than own everything from web-hosting resources to lab space and equipment. Nick Denton, founder of Gawker Media, a new-model publishing mini-empire in New York based on blogs, touts steps to tie employee wages to advertising revenue and page views — rather than having fixed salaries — as part of his recessionary strategy.

Although times are tough and getting tougher, today's start-ups have at least one big advantage over their predecessors. Decades of entrepreneurship have created a rich infrastructure to support new ventures. In addition to free software, low-cost web-hosting services and start-up-friendly lawyers and accountants, there is a deep pool of students and young people willing and eager to trade freshly minted academic knowledge for some entrepreneurial experience and 'sweat equity'. The net result is that starting and running a company can be less expensive than ever before.

Paul Graham, a successful entrepreneur who now mentors start-ups through the venture firm Y Combinator in Mountain View, California, says some of his advisee companies have become profitable with revenues of as little as \$3,000 a month. True, such companies are going to have to grow if they are to provide full-time support for founders, let alone any employees. But with its own profits to re-invest in growth, companies are free of the need for further finance; they can chart their own course into the future. In tough times, that is the key to success. ■

**John Browning** is a writer and consultant. He co-founded the First Tuesday networking forums for venture capitalists and innovators. e-mail: jb@poyais.com

We are witnessing the coming home to roost of the era of 'Gucci capitalism'. An era, born of the mid-1980s, that failed to question the market's ability to self-regulate and tilted the balance of power between corporations and society firmly in the favour of businesses. An era that deified self-interest and greed. It precipitated a global recession that will last until at least mid-2010, and triggered a new socio-political environment that presents a very different playing field for business. We are now in the era of 'co-op capitalism', where cooperation rather than selfish behaviour will rule.

To understand co-op capitalism requires an understanding of socioeconomic, political, psychological, historical and legal trends. And so it will bring work of often-sidelined researchers into the mainstream. Traditional business forecasts have usually ignored such 'soft' knowledge. But with the entire political and economic system in a state of flux, companies would be well-minded to monitor Internet campaigns, track media concerns, and search for historical parallels. When an economic trajectory is constant, such data might be considered incidental. When we don't even know what trajectory we are on it becomes essential.

Relying on information about interest rates to predict consumption, for example, is useless if fear and uncertainty drive people to buck traditional responses to low interest rates by saving rather than spending — as seems to be the case now. But history can give us clues as to what society wants. During past economic crises such

as the Great Depression, the public assumed more communitarian values. Historian Robert McElvaine of Millsaps College, Jackson, Mississippi, describes it as "the time in which the values of compassion, sharing and social justice became [more] dominant than they had ever been before in American history"<sup>1</sup>. This chimes with anthropological studies that point to less material wealth leading to a greater propensity among people to share what they have<sup>2</sup>. While these values have been largely eroded over the past twenty years, in the co-op capitalism era they are likely to re-emerge dominant. What is not clear is where the limits of cooperation will lie. Will our sense of community stretch beyond local and national borders, and reach others less fortunate in far-flung regions of the world?

Regardless, collective interest should become a guiding corporate principle. This means that investments in the community, volunteering opportunities for staff, and the provision of affordable products to the poor must now be recognized as integral parts of business strategy. Corporations should preempt any interventions on executive pay and come up with policies acceptable to their shareholders, investors, the government and the public first, and to their own executives second.

This pre-emptive strike is advisable, as public desire for state intervention is at a new record high. It is not only banks that are caught in the fray. Any industry perceived to be acting against the public good is likely to find itself in the line of fire. Acting selfishly will be penalized.

Take the fast-food industry. Michael Bloomberg, the mayor of New York, was ahead of the game when, at the end of 2006, he brought in a ban on the use of artificial trans-fats — substances known to increase obesity and that have suitable, if more expensive, alternatives. The United Kingdom has restricted the television marketing of 'junk' food to children. Anticipate more measures on that front, and similar targeting of other sectors too. Big pharmaceutical companies, for example, should be on the look out. With health costs soaring and governments needing to rein in expenditure, predict





more pressure on pharmaceutical companies to deliver affordable medicines.

The legacy of a corporate leader will in part be determined by whether she is perceived as being with or against 'us', the public — whether she shops at the Co-op or at Gucci. Choose the former for longer-term viability. ■

**Noreena Hertz** is a visiting professor of globalization at the Rotterdam School of Management, Erasmus University, the Netherlands, and a fellow of the Judge Business School, University of Cambridge, UK. Her most recent book is *IOU: The Debt Threat* (2005). e-mail: n.hertz@jbs.cam.ac.uk

1. McElvaine, R. S. *The Great Depression: America 1929–1941* (Three Rivers Press, 1993).
2. Mauss, M. *The Gift: Forms and Functions of Exchange in Archaic Societies* (Routledge & Kegan Paul, 1969).

## End the obsession with interest

Regulating leverage, not interest rates, is the answer to a troubled economy, says **John Geanakoplos**.

In standard economic theory, the interest rate has long been regarded as the most important variable. Whenever the economy slows, and asset prices fall, economists clamour for lower interest rates to encourage more spending, and the US Federal Reserve usually obliges. It has recently obliged again, lowering the bank rate to nearly zero. But sometimes, especially in times of crisis, it's the collateral a borrower needs to post (or what economists call leverage) that is far more important.

Shakespeare got this right 400 years ago. In *The Merchant of Venice*, when Shylock grants a loan to Antonio they negotiate not just the interest rate but a pound of flesh as collateral too. It is clear which Shakespeare thought was more important: nobody who sees the play ever remembers the interest rate, which was zero.

It is apparent from everyday life that the laws of supply and demand can determine both the interest rate and leverage of a loan: the more impatient borrowers are, the higher the interest rate; the more nervous the lenders become, the higher the collateral they demand. But standard economic theory fails to properly capture these effects, struggling to see how a single

supply-equals-demand equation for a loan could determine two variables: the interest rate and the leverage. The theory typically ignores the possibility of default (and thus the need for collateral), or else fixes the leverage as a constant, allowing the equation to predict the interest rate.

Yet variation in leverage has a huge impact on the price of assets, contributing to economic bubbles and busts. This is because for many assets there is a class of buyer for whom the asset is more valuable than it is for the rest of the public (standard economic theory, in contrast, assumes that asset prices reflect some fundamental value). These buyers are willing to pay more, perhaps because they are more sophisticated and know better how to hedge their exposure to the assets, or they are more risk-tolerant, or they simply like the assets more. If they can get their hands on more money through more highly leveraged borrowing (that is, getting a loan with less collateral), they will spend it on the assets and drive those prices up.

In the absence of intervention, leverage becomes too high in boom times, and too low in bad times. As a result, in boom times asset prices are too high, and in crisis times they are too low. This is the leverage cycle.

Leverage dramatically increased in the United States from 1999 to 2006. A bank that in 2006 wanted to buy a AAA-rated mortgage security could borrow 98.4% of the purchase price, using the security as collateral, and pay only 1.6% in cash. The average leverage in 2006 across all of the US\$2.5 trillion of 'toxic' mortgage securities was about 16 to 1, meaning that the buyers paid down only \$150 billion and borrowed the other \$2.35 trillion. Home-buyers could get a mortgage leveraged 20 to 1: a 5% deposit. Security and house prices soared.

Today leverage has been drastically curtailed by nervous lenders wanting more collateral for every dollar loaned. Those toxic mortgage securities are now leveraged on average only about 2 to 1. Home buyers can now leverage themselves at only 5 to 1 if they can get a government loan, and less if they need a private loan. De-leveraging is the main reason the prices of both securities and homes are still falling.

The leverage cycle is a recurring phenomenon. The financial derivatives crisis in 1994 that bankrupted Orange County in California was the tail end of a leverage cycle. So was the 1998 emerging markets/mortgage crisis.

I have spent the past ten years mapping out the leverage cycle, and although some aspects of it are now well understood, others are not. What signs reveal when the crisis is about to begin, and how best should the government

intervene? I am now working on these and other questions, in part under the auspices of the Santa Fe Institute in New Mexico with physicist Doyne Farmer.

It is clear that the leverage cycle cannot be stopped by blaming or shaming greedy Wall Street investors or by exhortations not to panic. The cycle emerges even if — in fact precisely because — every agent is acting rationally. But this individual rationality leads to collective disaster. The government must intervene.

What the Federal Reserve should do is manage leverage, curtailing it in ebullient times and propping it up in anxious times — especially in a crisis like now. Instead, it remains obsessed with managing the economy by lending money to banks at lower and lower interest rates, hoping, for no good reason, that the banks will turn around and lower the collateral requirements they impose on borrowers.

The upshot of *The Merchant of Venice* is that the regulatory authority (the court) intervenes, not by changing the interest rate, but by imposing a collateral level different from what Shylock and Antonio freely negotiated: a pound of flesh, but not a drop of blood. The Federal Reserve should follow Shakespeare's lead. ■

**John Geanakoplos** is the James Tobin Professor of Economics at Yale University, an external professor at the Santa Fe Institute and a founding partner at Ellington Capital Management, a hedge fund specializing in mortgages. email: john.geanakoplos@yale.edu

See also page 942.

# BOOKS & ARTS

## Old lessons for a new economics

Nobel prizewinner Paul Krugman's updated analysis of past economic crises teaches us that recovery now will require more than a new set of rules, explains **Bill Emmott**.



### The Return of Depression Economics and the Crisis of 2008

by Paul Krugman

Penguin/W. W. Norton: 2008.

256 pp/224 pp. £9.99 (pbk)/\$24.95 (hbk)

An often-told story about a politician — whose name varies according to the nationality of the teller — has him declaiming loudly that he is a man who sticks to his principles. But, he adds after a coy hesitation, if you don't like those principles, I have some others I could try.

That, essentially, is what Paul Krugman, a professor of economics at Princeton University and the 2008 Nobel prizewinner for his discipline, says is happening to economics. Rules that had become commandments — thou shalt not run a big budget deficit, nor allow the money supply to grow excessively, nor take private firms into public ownership — are being thrown away amid the worst economic downturn since the early 1980s, and possibly since 1945.

Krugman does not decry this abandonment of principles. On the contrary, he argues that it hasn't happened fast enough. For these were not laws of economics but rather creatures of circumstance that were misunderstood or misused. For more than three decades, the world's industrialized economies suffered from inflation and were preoccupied by trying to control or even defeat it. High public borrowing and rapid monetary expansion are dangerous only when they cause or contribute to inflation, which they did during the 1970s and 1980s, by adding further to demand at a time when too much money was already chasing too few goods and services.

That is not, however, the situation now. Rather than excessive demand and inflexible supply, developed economies are facing inadequate demand. That problem occurred during the Great Depression of the 1930s, and it is what Krugman means by 'depression economics'.

It is the topic that made John Maynard Keynes the most famous economist of the 1930s, with his book *The General Theory of Employment, Interest and Money* (1936). This work provided the theoretical framework for the use of public borrowing and spending to counter the insufficiency of private demand during depressions, although the practice of that policy in the United States, in the form of President Franklin D. Roosevelt's New Deal, was already under way by then. The Keynesian idea was that fiscal policy could be

remarks, which some say is apocryphal, was made when he was challenged over an apparent change of position. He is supposed to have said: "When the facts change, I change my mind. What do you do, sir?" That point is made here by Krugman, a long-time fan of Keynes. Minds should change when the facts change. Facts did change during several of the financial crises of the 1990s — in Japan, Mexico, east Asia, Brazil and elsewhere — when depressionary tendencies were seen, yet minds still stuck wrongly to the old orthodoxies.

This rigidity of thinking, when combined with a desire to believe that deregulated financial markets would be capable of learning from their own mistakes, has led to the current global economic crisis.

*The Return of Depression Economics and the Crisis of 2008* is a fine piece of popular writing. Krugman avoids jargon and technicalities without leaving the already-informed reader either bored or annoyed, using homely examples to describe economic issues. For example, he makes good use of a story of how a babysitting cooperative on Washington's Capitol Hill created its own depression. The cooperative worked by giving coupons to those who babysat, which they could then use to buy babysitting services from other members; the depression occurred when too many members began to hoard coupons for later use, so that demand fell well below supply.

That, more or less, is what is happening now. Shocked by the collapse of banks on both sides of the Atlantic, households and companies are increasing their savings and paying off their debts to protect themselves against harder economic times. Individually sensible, such action to save more

is collectively disastrous: what Keynes called "the paradox of thrift". Demand is slumping, production lines are being halted, people are losing their jobs and so demand is likely to fall further. At such times when people are scared of debt, they will not borrow even if the cost of money falls virtually to zero: a 'liquidity trap'.



Old tricks: President Roosevelt saw spending as the economy's saviour in 1938.

used to fine-tune the economy, ensuring that there would be full employment at all times. It was the misuse of that notion in the 1970s that helped to entrench inflation, which in turn made the need to restrict budget deficits seem like a fixed principle.

One of Keynes's most famous off-the-cuff

What to do? The weakness of this book, an update of a work first published in 1999, is that its emphasis is on the financial crises of the 1990s and not that of today. If Krugman were writing from scratch now, he would surely focus much more on the present day. However, the 11 pages he does devote to solving the current crisis make good sense: government borrowing is vital to substitute for slumping private demand, as Keynes said. At the same time, banks need to be recapitalized and have their bad debts written off to make them capable of lending again; financial regulation must

be reformed to make the system more stable and to encourage people to trust it again. All that is fine in principle. But more detail would have helped readers to understand how it might work in practice. ■

**Bill Emmott** was editor of *The Economist* during 1993–2006, and is the author of *Rivals: How the Power Struggle between China, India and Japan will Shape our Next Decade*.

e-mail: bill@billemott.com

For more on the credit crunch, see online at [www.nature.com/news/specials/recessionwatch](http://www.nature.com/news/specials/recessionwatch).

## Investing in the environment

### Sustainable Investing: The Art of Long-Term Performance

Edited by Cary Krosinsky and Nick Robins  
Earthscan: 2008. 272 pp. £19.99, \$38.95

Money can't buy love. But can it buy a more sustainable world? *Sustainable Investing* suggests that it can, yet this book presents only half of the story.

Editors Cary Krosinsky and Nick Robins are industry people: Krosinsky is vice-president of a US environmental research organization and advised on the United Nations' Principles for Responsible Investment as part of its Environment Programme Finance Initiative, and Robins is head of the HSBC Climate Change Centre of Excellence. Together with co-authors from a range of financial organizations, they provide tips on the emerging market of sustainable investing, ranging from equities, bonds, carbon finance, private equity, sustainable real-estate and microfinance. With up to US\$5 trillion in sustainable investment funds and a quarter of public equities and bonds incorporating sustainability criteria, the world has never had such green markets.

If you take this book at face value, it certainly pays to be green. Sustainable investing — investing in green technologies, clean-energy sectors or companies that have an integrated environmental approach — has “outperformed not only ethical funds, but mainstream indices as well over a one-, three- and five-year period”. Ethical funds, a separate category of responsible investing, are less profitable.

You may be relieved to hear that your clean-energy funds are good

financial bets. But Krosinsky and Robins warn us against the ‘materiality trap’: the tendency of investors to incorporate environmental or social issues only when they are financially attractive. Many chapters focus on this point. Yet the book does not pay enough attention to another trillion-dollar question: how do we measure the effects of investment on the sustainability of our world?

To answer that, we need to know more about the systemic effects of specific investments on specific ecosystem services and societies, both over time and on a cumulative basis. This requires multidisciplinary research. For example, funds such as Jupiter Ecology, Winslow Green Growth, Orange SeNSE or the CalPERS emerging markets fund all invest in green activities and offer strong financial returns. It is less clear if they make measurable

reductions in the degradation of various ecosystems. I am not saying they do not — it is just very difficult to tell.

In the financial districts of New York, Hong Kong and the City in London, market intelligence is not set up to convey this kind of knowledge. Research shows that capital markets have difficulty capturing environmental feedback. Most financial-investment decisions are decoupled from on-the-ground effects. The abstraction becomes clear only when there is a visible collapse.

In one chapter, investment managers from the investment banking group Société Générale concur: “In our experience, long-term analysis is almost always sacrificed in favour of short-term profits by both investors and brokers.” Sustainable investment can introduce new information, but there is significant bounded rationality. In regular financial investment, “a financial analyst covers a maximum of four to eight stocks and often has as much as 10 or 20 years’ experience in doing so. In contrast, many SRI [socially responsible investment] analysts cover up to 50 stocks, with far less experience and an often limited financial background.” There is also the inherent danger of market spin in sustainability ratings: “best in class … needs to be complemented by a searching analysis of the sustainability of the ‘class’ itself”. The authors warn that unless this is done, a situation can emerge in which everyone wins, and relative benchmarking ensures that there is always a sector leader, even if neither business model nor behaviour warrant such an outcome.

*Sustainable Investing* argues for a paradigm shift in financial markets, and suggests that better data and stronger networks will positively affect the effectiveness of managers of sustainable funds. However, this is not an academic book, and the chapters are of uneven quality. Notably absent is any mention of what role scientists could have in this shift — a serious omission.

From a scientific standpoint, managing complex nonlinear change requires flows of robust information that cross spatial and temporal scales, and adaptive governance arising from a mix of legislation and economic incentives. Sustainable investing is only one of many strategies that we can choose, and it should be more tightly coupled with actual impacts.

The recent financial collapse highlighted the need for new forms of financial governance. To take advantage of this opportunity, we need to ask questions that are more



Green technologies, such as this reverse vending machine that refunds consumers who recycle, represent good investments.

fundamental. For example, how can dynamic socio-ecological feedback loops become integrated within dynamic financial markets? How can we measure the effects of trillions of dollars of investment on natural capital stocks or ecosystem services? How can financial governance models make these interconnections possible? And we need to experiment with new organizations that better connect financial markets to science and societies.

One such example is the Resilience Alliance, an international research network of scientists and practitioners working on sustainability. Their Connectors Group is trying to establish more-explicit links with the business world — not an easy task given the differences in career

focus, social norms and conventional meeting places. Yet scientists can be inspired by the chapter on civil society. Non-governmental organizations influence capital markets through active capital campaigns, including shareholder activism, partnerships with powerful institutional investors and formal investor briefings.

*Sustainable Investing* is a good read for managing your portfolio. But it remains to be seen if *Nature* will be read by investment analysts. ■

**Gail Whiteman** is associate professor and director of the Sustainability and Climate Research Centre, Rotterdam School of Management, Erasmus University, 3062 PA Rotterdam, the Netherlands. e-mail: gwhiteman@rsm.nl

## In search of adventure

Most scientists have tried to explain their work to inquisitive relatives or acquaintances. Whether one studies salamander biology, glacial moraines or dark energy, one inevitably faces someone who is convinced that a life devoted to understanding nature is no vocation for a grown person and is merely an escape from the 'real world'. The string theorist and author, Brian Greene, has argued that these misperceptions are derived from how science is taught and communicated to the public. His prescription is that scientists need to convey the breathtaking vistas opened up by scientific inquiry and teach science for what it is: one of humanity's great adventure stories. Evolutionary biologist Sean Carroll's new book, *Remarkable Creatures*, fills this important niche.

Carroll tells the story of evolution through the adventures that lie behind its great discoveries. With skill as a storyteller, and a passion for his subject, Carroll deftly describes the great expeditions and people behind 13 major discoveries in natural history over the past 200 years. He imparts an infectious enthusiasm for the act of discovery, with all the risks, failures, dead ends and ultimate successes that expeditionary work often involves. As fast-paced as a detective story, the book is a primer on the joy of being an evolutionary biologist, written by one of our best modern practitioners.

Carroll reveals the challenges and rewards of venturing into the unknown. He begins his tour in 1800, in an age when vast stretches of



Ecuador's unexplored terrain thrilled the nineteenth-century naturalist and explorer Alexander von Humboldt.

### Remarkable Creatures: Epic Adventures in the Search for the Origin of Species

by Sean B. Carroll

Houghton Mifflin Harcourt: 2009.

352 pp. \$26

Earth were unexplored wilderness, and much of the flora and fauna was unknown to western science. His first tale, that of the naturalist and explorer Alexander von Humboldt, opens with the combination of enthusiasm, hardship and adventure that pervades the narrative of

later chapters. Venturing into the forests of South America, Humboldt exclaims in a letter to his brother: "What trees! Coconut trees, fifty to sixty feet

high... We rush around like the demented; in the first three days we were unable to classify anything..." Like Humboldt, each of the scientists in Carroll's book was the first one to see a new fossil, species or historical connection. And each shared Humboldt's excitement.

The first part of the book describes the travel

adventures of Charles Darwin, Alfred Russel Wallace and Henry Walter Bates. Lacing each tale with anecdotes of success and failure, Carroll conveys the sense that real people with real aspirations, fears and enthusiasm endured years in jungles and on ships to come up with ideas that would ultimately change the way scientists view the natural world. Next come stories of major fossil discoveries, showing how exploring the rocks of the world can reveal the history of life. He uses as his examples Eugène Dubois's discovery of the Java Man fossils, the Cambrian radiation of animal phyla, the origin of birds, the impact theory for mass extinction, and the fishapod, *Tiktaalik*, including my own work with Ted Daeschler and Farish A. Jenkins. The final grouping looks at human origins and bridges from the fossil record to the molecular one, first revealed by Linus Pauling and Emile Zuckerkandl, and later by Allan Wilson, Vincent Sarich, Mark Stoneking and Svante Pääbo. This shift, from expeditions revealing new fossils to laboratory adventures recovering ancient DNA, reflects a bridge to our current approach, one that integrates data from fields as diverse as geology and molecular biology to explain evolutionary history.

In reading *Remarkable Creatures*, one could feel a sense of loss that the days of great adventure are no more. We live in an age in which almost every region of Earth has been mapped and every species known is being documented with its own web page in the Encyclopedia of Life. It is easy to think that the days when a naturalist such as Humboldt could describe themselves as demented with the joy of seeing a new vista for the first time are over.

But unexplored vistas still exist, from the countless new species of microbes that dwell in every clump of dirt and every inch of our guts to the deepest reaches of space revealed by telescopes and probes. Indeed, in the cast of scientists manning the Mars rovers at NASA mission control, we find a room full of Humboldts, exploring a new world and giddy with each patch of Mars that is seen for the first time. These emotions are among the universals of our vocation, and there are few better introductions to them than Carroll's book. *Remarkable Creatures* is the book to give those inquisitive relatives, or even more important for our society's future, their kids. ■

**Neil Shubin** is professor of anatomy at the University of Chicago and author of *Your Inner Fish: A Journey into the 3.5 Billion Year History of the Human Body*.  
e-mail: nshubin@uchicago.edu

# Q&A: Building on paradise

Communicating the ideas of evolution is as much a challenge now as it was 150 years ago. In the wake of his recent BBC television programme, *Charles Darwin and the Tree of Life*, naturalist and broadcaster **David Attenborough** tackles those who challenge evolution head on.



Darwin200

## Why is teaching evolution now more important than ever?

Because of the influence of the Bible's book of Genesis, which says the Lord God said 'go forth and multiply' to Adam and Eve and 'the natural world is there for you to dominate, you have dominion of the animals and plants of the world'. That basic notion — that the world is there for us, and if it doesn't serve our purposes it's dispensable — has produced the devastation of vast areas. We have assumed that we can build a house on it, dig it up, put tarmac over it; that's OK because it's there for us. In finding solutions to our ecological problems we have to understand evolutionary processes.

## What are you trying to convey in *Charles Darwin and the Tree of Life*?

It's not a programme about Darwin's life. It's about the objections that people in Darwin's time and since have found to avoid accepting the propositions of evolution. In Darwin's time the geographical distribution of animals was not understood, the links between the great phyla were believed to be missing, the age of Earth was thought to be too short to have allowed evolution to take place, and the mechanism of genetics was not understood at all. The answers to those questions have only emerged in my lifetime. We've been able to precisely date rocks because of radioactivity for only 50 years; we have understood continental drift for 50 years and DNA for less time than that. But cumulatively those answers validate Darwin's insight.

## Do you think television is a good medium for science?

It's a very good medium for lighting flames of curiosity, raising questions, getting people excited. I don't think it's a very good medium for working things out in fine detail. People who are interested in finding out more go to books. Books, or writings on your computer screen, which you can take at your own pace, are much better for explaining scientific theory. But television creates excitement.

## How do you interact with scientists when making your programmes?

I say 'How I envy you, in my romantic way, having drawn the veil away from the face



One of David Attenborough's favourite moments was watching the vibrant bird of paradise (right).

of truth and seen a little bit of something nobody else has ever seen — how wonderful that must be! And they say, 'Oh yeah, but it's taken me 35 years of sitting in this swamp to do this; I wish I had the opportunity to go to the desert or the poles or wherever you have been.' So you can't have everything.

## After 50 years of filming the natural world, what is your favourite moment?

Watching displaying birds of paradise, because they are very remote from us. When you look at birds of paradise you realize there's an essence of life — an extraordinary life force — that has nothing to do with humanity and that has been going on long before primates even appeared on Earth. When, sitting concealed in a hide, you see all that vigour and variety and passion, you realize that humanity is not dominant in the world and is only a small part of it. I find that moving.



## What next?

I'm going to South Georgia in the south Atlantic to look at the glacier that I visited 15 years ago, which has now virtually disappeared. The influence of humanity on the natural world has become much greater in the past decade. That's why it's important to tell people about evolution. Still in the back of so many minds is the thought that 'we own the world and we can do what we like with it, and if we wish to devastate it that's entirely up to us.' It is not entirely up to us.

Interview by **Adam Rutherford**, Nature's Podcast and Video editor.

Watch the interview on Nature's YouTube channel at [www.youtube.com/NatureVideoChannel](http://www.youtube.com/NatureVideoChannel).

## NEWS &amp; VIEWS

## CARBON CYCLE

# Sink in the African jungle

Helene C. Muller-Landau

**Apparently pristine African tropical forests are increasing in tree biomass, making them net absorbers of carbon dioxide. Is this a sign of atmospheric change, or of recovery from past trauma?**

The lush vegetation of tropical forests is a large and globally significant store of carbon<sup>1</sup>. Because tropical forests contain more carbon per unit area than any alternative land cover, cutting them down releases carbon into the atmosphere. For the same reason, growing forests take up carbon from the atmosphere. Of course, trees cannot grow for ever, and neither can forests: in the absence of disturbances that kill trees en masse — such as fires, hurricanes or logging — every forest will eventually reach a point at which tree growth and death are in equilibrium, and at which the average change in tree carbon stocks is zero.

It is thus surprising that undisturbed tropical forests currently do not seem to be at equilibrium. If you measure the size of trees in a given area, calculate their carbon stocks, and then repeat the process some years later, you will on average find that the forest holds more carbon than it did before. This was first reported for Amazonian tropical forests<sup>2</sup>, and on page 1003 of this issue Lewis *et al.*<sup>3</sup> show that African forests also have increasing stocks of tree carbon.

So how much carbon are we talking about? Using data collected in Africa between 1968 and 2007, the authors find that trees have added an average of 0.63 tonnes of carbon per hectare each year. Given that approximately half the dry matter in trees is carbon, the amount of wood added annually in each hectare of African forest is equivalent in mass to a small car. For comparison, the average rate of carbon accumulation in tropical forests around the globe was 0.49 tonnes of carbon per hectare per year<sup>2–4</sup>. Extrapolating from their data<sup>3</sup> by assuming parallel changes in the carbon pools of roots and dead trees, Lewis *et al.* estimate that 'old-growth' tropical forests are taking up  $1.3 \times 10^9$  tonnes of carbon per year worldwide.

There are two possible explanations for this finding. One is that the tropical forests that



**Figure 1 | Getting bigger.** Lewis *et al.*<sup>3</sup> show that apparently undisturbed African tropical forests are currently increasing in tree biomass each year, and act as carbon sinks. But it is impossible to say how long this will continue.

we think of as intact actually suffered major disturbances in the not-too-distant past, and are still in the process of growing back<sup>5</sup>. This recovery process is known as succession, and takes hundreds — or even thousands — of years. Succession involves not only initial growth to full canopy height, but also subsequent gradual shifts in species composition. The past disturbances could have been natural or anthropogenic; possible explanations include droughts and fires related to huge El Niño events, and changes in land use that

allowed previously cleared land to revert to forest<sup>5</sup>.

In fact, palaeoecological and archaeological evidence increasingly documents the long disturbance histories of today's 'undisturbed' tropical forests<sup>6</sup>. There have been many large fires in Amazonian forests over the past few millennia, the timings of which are related to both climate and the size of human populations<sup>7</sup>. Far from being pristine wildernesses little influenced by their human inhabitants, many areas were cleared or otherwise intensively used in centuries past<sup>8</sup>. Given the timescales of tropical-forest succession, these disturbances are almost certainly contributing to carbon accumulation in many tropical forests today.

The second explanation for Lewis and colleagues' findings<sup>3</sup> is that tropical forests have been knocked from their previous equilibrium by global climate and/or atmospheric change<sup>9</sup>, so that they are currently in transition to a higher carbon state. Perhaps, for example, the increase in atmospheric carbon dioxide is effectively fertilizing tropical tree growth. Under these circumstances, if tree mortality doesn't keep pace with increases in growth, then trees will on average grow larger before they die (Fig. 1), and tree carbon stocks will increase<sup>10</sup>. Carbon stocks in mature tropical forests vary enormously depending on climate, soil type and topography; temporal changes in climate and resource availability would therefore be expected to have parallel influences in the long run.

The two mechanisms that might account for Lewis and colleagues' observations<sup>3</sup> would be expected to produce different spatial and temporal patterns of carbon uptake by trees, but our current knowledge does not allow us to predict what these patterns are, or to say which mechanism is operating in Africa. Over the course of succession, tree carbon stocks increase at an ever-slower rate as stands age. Thus, we

expect tree carbon stocks and their rate of change, and stand age, to be closely related within any given forest type. In temperate and boreal forests, where stand age is generally well known, carbon stocks and fluxes do indeed show a strong relationship with stand age, even at ages many consider to be old-growth<sup>11</sup>. Examination of these relationships in tropical forests is stymied not only by lack of information about how long ago disturbances occurred, but also by limited knowledge of how growth rates and equilibrium carbon stocks are affected by rainfall, soils and other factors.

One might suppose that predictions based on the global-change hypotheses are more straightforward — after all, atmospheric CO<sub>2</sub> concentrations are rising equally everywhere. In fact, the effects of CO<sub>2</sub> fertilization on tree growth are expected to depend strongly on other factors that vary greatly among forests, especially the availability of soil nutrients<sup>1</sup>. And if changing climate (rather than rising atmospheric CO<sub>2</sub>) is affecting the carbon flux of tropical forests, then the outcomes will differ depending on local changes and the local baseline.

It is likely that both succession and global change have a role in explaining tropical-forest growth, with varying importance at different sites. Where recovery from disturbance drives tropical-forest change, associated changes in species composition would be expected. A study published last year<sup>4</sup> found that tree species with slower growth rates are disproportionately increasing in biomass in nine out of ten 'undisturbed' tropical forests around the globe, as would be expected during succession. Yet Lewis *et al.*<sup>3</sup> find no relationship between a species' wood density and the rate of change of its population across their African plots.

A better understanding of tropical-forest carbon dynamics is clearly needed to determine the causes of the observed increases in tropical tree carbon stocks — and, more critically, to predict the future trajectory of these stocks under global change. Furthermore, we must look not only at the trees<sup>3</sup>, but also at the soil: tropical-forest soils hold at least as much carbon as the trees. Unlike tree carbon stocks, soil carbon stocks can potentially increase indefinitely. But the prevailing prediction is that increasing temperatures will speed decomposition and reduce soil carbon stocks.

In the future, will tropical trees and soils act as carbon sinks, thereby slowing atmospheric and climate change? Or will altered climates turn them into carbon sources that accelerate further change? Standardized assessments of the main carbon pools and fluxes of tropical forests around the world — and their proposed drivers — are needed to document and understand the current trends, to inform predictive models, and ultimately to answer these pressing questions.

Helene C. Muller-Landau is at the Smithsonian Tropical Research Institute, Apartado Postal

0843-03092, Balboa, Ancon, Panamá, República de Panamá.  
e-mail: mullerh@si.edu

1. Solomon, S. D. *et al.* *Climate Change 2007: The Physical Science Basis* (Cambridge Univ. Press, 2007).
2. Phillips, O. L. *et al.* *Science* **282**, 439–442 (1998).
3. Lewis, S. L. *et al.* *Nature* **457**, 1003–1006 (2009).
4. Chave, J. *et al.* *PLoS Biol.* **6**, e45 (2008).

5. Wright, S. J. *Trends Ecol. Evol.* **20**, 553–560 (2005).
6. Clark, D. A. *Biotropica* **39**, 4–19 (2007).
7. Bush, M. B., Silman, M. R., McMichael, C. & Saatchi, S. *Phil. Trans. R. Soc. Lond. B* **363**, 1795–1802 (2008).
8. Mann, C. C. *Science* **321**, 1148–1152 (2008).
9. Lewis, S. L. *Phil. Trans. R. Soc. Lond. B* **361**, 195–210 (2006).
10. Muller-Landau, H. C. *et al.* *Ecol. Lett.* **9**, 589–602 (2006).
11. Luyssaert, S. *et al.* *Nature* **455**, 213–215 (2008).

## NEUROSCIENCE

# Good and bad cell death

Donald W. Nicholson

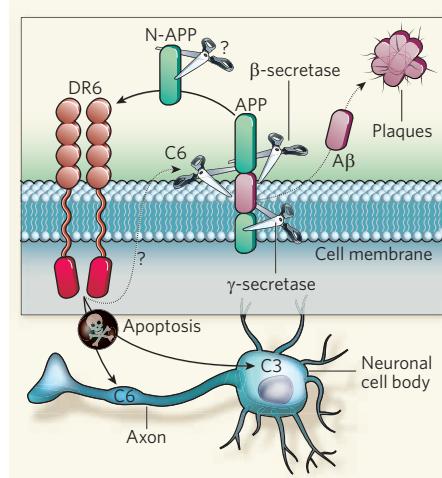
**Neurodegeneration often has disease connotations. However, it is also a developmental process for fine sculpting of the nervous system. One signalling cascade might mediate the process in both circumstances.**

A major driver of neurodegeneration in Alzheimer's disease is thought to be plaques of amyloid-β peptides. These 40–42-amino-acid fragments form when the transmembrane amyloid-β precursor protein (APP) is degraded by the enzymes β- and γ-secretase. In addition to amyloid-β peptides, degradation of APP results in the formation of other peripheral fragments<sup>1</sup>. Perhaps unsurprisingly, given the compelling link between the deposition of amyloid-β peptides and the development of Alzheimer's disease<sup>2,3</sup>, these additional fragments have been outshone and have received scant attention. But on page 981 of this issue, Nikolaev and colleagues<sup>4</sup> report that one oft-neglected fragment — the soluble amino-terminal portion of APP (N-APP) — has a crucial role in fine-sculpting of the nervous system during development, and possibly in disease-associated neurodegeneration as well.

In the nervous system, cell death is a surprisingly common and immensely robust process, especially during development, when productive neural circuits are reinforced and non-productive ones are eliminated. The neurons that fail to establish appropriate links with other neurons are culled mainly by apoptosis, a process that has the cell 'commit suicide' and then package up its residual bits for recycling. Those neurons that forge essential connections are protected from apoptosis by 'trophic' proteins such as nerve growth factor (NGF).

To engage other cells, neurons sprout axonal extensions, which, when necessary, are also pruned by similar, although not identical, mechanisms. So, just as Michelangelo described his technique for the creation of masterpieces such as the *Pietà* saying, "Every block of stone has a statue inside it and it is the task of the sculptor to discover it", the intricate biochemical processes of biogenesis and death determine the formation, maintenance and capabilities of the nervous system.

One approach by which to study these



**Figure 1 | Neurodegeneration signalling cascade.** The transmembrane amyloid-β precursor protein (APP) is routinely cleaved by β- and γ-secretase enzymes to generate both amyloid-β peptide (Aβ) and the amino-terminal portion of APP (N-APP). The soluble N-APP may then undergo further processing before binding to the death receptor DR6. Nikolaev *et al.*<sup>4</sup> find that DR6 responds to interaction with N-APP by engaging and activating key mediators of apoptotic cell death: caspase-3 (C3) in the neuronal cell body and caspase-6 in axons. Caspase-6 (C6) can also cleave APP near the β-secretase target site, potentially contributing to the formation and/or amplification of this apoptotic circuit.

intricate, sophisticated and little-understood processes is to grow neurons in culture and then induce apoptosis by withdrawing NGF. Using this method in three types of spinal-cord neuron, Nikolaev *et al.*<sup>4</sup> find that the apoptotic machinery of these cells responds to NGF withdrawal as expected — neurons commit suicide or axons degenerate. In each case, however, initiation of apoptosis seems to depend on the presence of the 'death receptor' DR6.

Members of the death-receptor family sense

extracellular apoptotic signals — death ligands — on the cell surface and bind to them to trigger apoptosis<sup>5,6</sup>. The prototypical death receptors are Fas/CD95 and tumour-necrosis factor receptor 1 (TNFR1); once bound to their ligand, these receptors sequester components of the cell-death pathway — such as inactive precursors of the caspase enzymes — within the cytoplasm, thus activating them<sup>7</sup>. Each of the other known death receptors functions through subtle permutations of this process (different ligands, adapters and points of entry into the caspase cascade). In this regard, little information exists for DR6, which has thus been considered ‘orphan’.

Nikolaev *et al.*<sup>4</sup> show that, on withdrawal of NGF from their cultured neurons, N-APP activates DR6, which then triggers activation of caspase-3 in neuronal cell bodies. In the axons, however, caspase-6 — something of an orphan enzyme itself — is the primary responder to N-APP engagement of DR6.

The enigmatic caspase-6 is often classified as an effector caspase because it cleaves the nuclear lamin proteins during apoptosis. Nonetheless, its substrate preferences indicate that — similarly to caspases-8, -9 and -10 — it probably functions early in the apoptotic pathway, before ‘death’ signals reach the nucleus.

Strikingly, caspase-6 can itself liberate N-APP by cleaving APP in almost the same spot as the  $\beta$ -secretase<sup>8</sup> (Fig. 1). So, although Nikolaev *et al.*<sup>4</sup> provide evidence that  $\beta$ -secretase is the initial perpetrator of N-APP release from the cell surface, it is possible that the newly activated caspase-6 feeds back to liberate more N-APP, thus either amplifying the apoptotic process or spreading it to neighbouring cells. Such potential secondary effects are hard to ignore, particularly because they might be relevant to Alzheimer’s disease.

But disease-associated effects of caspase-6 are not limited to Alzheimer’s disease: this bad-boy caspase contributes to at least one other neurodegenerative disorder, Huntington’s disease<sup>9</sup>. Whereas Alzheimer’s disease affects neurons and synaptic junctions of the cerebral cortex, Huntington’s disease is characterized by progressive and inexorable deterioration of neurons that project to the striatum region of the brain. Caspase-6-mediated breakdown of huntingtin, the protein that is mutated in Huntington’s disease, is necessary for neuronal dysfunction and degeneration in this disorder. Whether the circuitry involved in APP cleavage, DR6 triggering and caspase activation have broad, overlapping mechanistic commonalities in the development of the nervous system, response to injury and disease-associated neurodegeneration is not known. But the links are intriguing and warrant further attention.

The molecular events that Nikolaev and colleagues describe might seem linear and binary (with loss of NGF resulting in APP cleavage to generate N-APP that binds to DR6 and activates caspases). But the reality is

probably that there is a dynamic equilibrium, and that subtle perturbations to that fine balance can evoke either necessary responses, including culling of neurons or pruning of their axons, or inappropriate ones, such as neurodegeneration. APP, for example, is continually generated and processed, forming a fluctuating pool of amyloid- $\beta$  peptides and N-APP. Not all amyloid- $\beta$  peptides form plaques, nor does N-APP always kill neurons. But as we come to understand the pathways and interrelationships within these sophisticated systems better, so will we come to understand what goes wrong and why. ■

Donald W. Nicholson is at Merck Research Laboratories, Rahway, New Jersey 07065, USA. e-mail: donald\_nicholson@merck.com

- Thinakaran, G. & Koo, E. H. *J. Biol. Chem.* **283**, 29615–29619 (2008).
- Mattson, M. P. *Nature* **430**, 631–639 (2004).
- Walsh, D. M. & Selkoe, D. J. *J. Neurochem.* **101**, 1172–1184 (2007).
- Nikolaev, A., McLaughlin, T., O’Leary, D. D. M. & Tessier-Lavigne, M. *Nature* **457**, 981–989 (2009).
- Pan, G. *et al.* *FEBS Lett.* **431**, 351–356 (1998).
- Ashkenazi, A. & Dixit, V. M. *Curr. Opin. Cell Biol.* **11**, 255–260 (1999).
- Kumar, S. *Cell Death Differ.* **14**, 32–43 (2007).
- Gervais, F. G. *et al.* *Cell* **97**, 395–406 (1999).
- Graham, R. K. *et al.* *Cell* **125**, 1179–1191 (2006).

## MEDICAL IMAGING

# MRI rides the wave

Paul Glover and Richard Bowtell

An innovative approach for exciting and detecting signals in magnetic resonance imaging not only improves image quality but also enables radical changes in scanner design by freeing up space around the patient.

Magnetic fields varying at radio frequency (RF) are fundamental to nuclear magnetic resonance (NMR) and the related technique of magnetic resonance imaging (MRI). In NMR, RF fields are used in conjunction with a strong, constant magnetic field to excite hydrogen nuclei in water into precession. The precessing nuclei in turn generate an RF magnetic-field oscillation — the NMR signal. In almost all MRI experiments, the RF field is generated by a coil of wire — the RF coil<sup>1</sup> — and the nuclear precession is detected by electromagnetic induction in a similar coil. But excitation and detection of the NMR signal require that the RF coils lie in close proximity to the human body because it is the short-range signals (the near fields) that are exploited in the interaction.

In a radical rethink of the experimental set-up used for MRI, Brunner *et al.*<sup>2</sup> (page 994 of this issue) now show that this conventional approach based on RF coils can be replaced by one that uses travelling radio waves in a long-range interaction with the sample, offering more-uniform coverage of larger samples. What’s more, by freeing up space in the bore of the scanner, this innovative approach could make the scanning experience more comfortable for human patients.

In a scanner with magnetic-field strength of 1.5 tesla (T) — the workhorse of clinical MRI — the RF required to excite the NMR signal is 64 megahertz (MHz), which corresponds to a free-space wavelength of 4.7 metres. This wavelength is reduced to about 70 centimetres in the human body as a result of the electrical properties (the permittivity and conductivity) of tissue. The near-field approach works well for this frequency regime because it readily produces a spatially uniform RF magnetic

field inside the body. This means that the signal excitation is homogeneous and, consequently, that magnetic resonance images show a uniform sensitivity to structure and pathology in different anatomical regions.

Magnetic resonance scanners operating at magnetic-field strengths higher than 1.5 T are increasingly being used because they provide higher sensitivity and spatial resolution, with 3 T (128 MHz) scanners becoming commonplace and scanners operating at 7 T (300 MHz) and above appearing in research environments. But at these operating fields the wavelength of the RF field is reduced and can become smaller than the structures of both the RF coil and the human body. As a result, standing waves formed by the RF coil give rise to significant spatial inhomogeneity of the RF amplitude, which degrades image quality at 3 T and can produce areas of total signal loss in images acquired at 7 T. Brunner *et al.*<sup>2</sup> show that these problems can be overcome by replacing the RF coils with a waveguide and a remote antenna. The new approach uses travelling waves for NMR signal excitation and detection, and its advantages result from the intrinsically lower spatial variation of RF amplitude in a propagating rather than standing wave pattern.

In their proof-of-principle experiments, Brunner *et al.* inserted a cylindrical waveguide into a 7 T human scanner and positioned an antenna at one of its ends (Fig. 1c on page 994). The waveguide diameter was just large enough to sustain a spatially uniform RF magnetic field at 300 MHz. A human patient sited inside the waveguide could thus be exposed to a homogeneous far-field wave pattern originating from the antenna. The same antenna was



## 50 YEARS AGO

A brief review of the history of the calendar by Y. G. Perel suggests the urgent desirability of establishing a world calendar, such as was proposed by India in 1953 before the United Nations. According to this proposal, the year will be divided into four quarters of thirteen weeks each, with the first month of thirty-one days and the following two of thirty days each. An additional day (the day of peace and friendship) is added after December 30, and on leap years an additional day after June 30.

From *Nature* 21 February 1959.

## 100 YEARS AGO

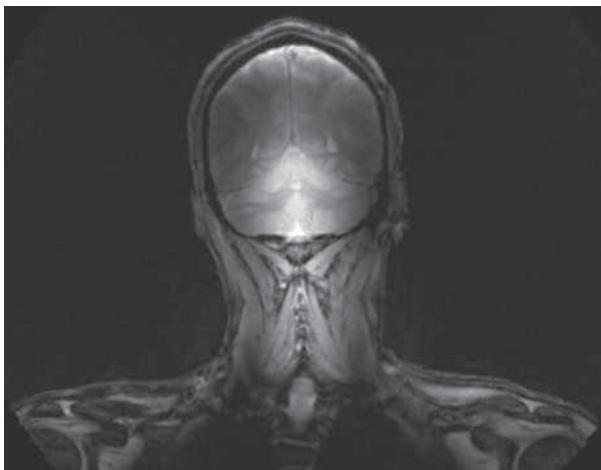
In this day of encyclopaedias numerous and ponderous, one is often struck with the fact that in spite of the manifest care and conscientious thought bestowed by the responsible editors, the omissions and evidences of discontinuity of treatment, and lack of recognition of the prime purposes of the compilation, are as noteworthy as the imposing array of the results of our steadily advancing knowledge is startling ... As an illustration, take the word "research," or any of the associated terms — "discovery," "experiment," "investigation," and "observation." Turning to the index volumes of the ninth and tenth editions of the "Encyclopaedia Britannica," I find but two references in which the word "research" appears — one to the exploring vessel, the *Research*, and the other to "research degrees."

### ALSO:

The *Petit Journal* recently asked its readers to select by their votes twelve great Frenchmen worthy of being included in the Pantheon. Pasteur's name appeared at the top of the poll with 315,203 votes, and was followed by that of Gambetta with 279,443 votes. We wonder whether a man of science would head the list if a similar plebiscite were taken by a popular daily paper in this country.

From *Nature* 18 February 1909.

50 & 100 YEARS AGO



**Figure 1 | Travelling-wave MRI.** Magnetic resonance image of the head and shoulders of a human volunteer obtained by Brunner *et al.*<sup>2</sup> using their travelling-wave MRI technique at a magnetic-field strength of 7 tesla. It demonstrates an impressive extent of coverage for an image obtained at such a high field: most conventional 7-tesla scanners are used for imaging only the brain, but this image shows detailed structure in the brain stem, neck and shoulders.

used to pick up the RF signal generated by nuclei in the body. Brunner and colleagues used this arrangement to record spectra and images from large test samples and the lower leg of a human volunteer. Figure 1 shows an additional image of the head and shoulders. Comparison with images obtained using a conventional RF coil shows that the travelling-wave approach can produce a uniform excitation over a much larger volume (Fig. 4 on page 996).

Until now, methods for moderating the effect of RF inhomogeneity at high magnetic-field strengths have generally focused on using multiple RF coils in parallel<sup>3</sup> for signal excitation. Although this approach shows promise, it is complex and expensive because it requires duplication of costly RF amplifiers and other circuitry. The beauty of Brunner and colleagues' travelling-wave approach, in comparison, lies in its simple implementation and inherently lower cost. In addition, the authors' approach has the advantage of being immediately compatible with all standard MRI techniques.

Although the main motivation of Brunner and colleagues' work is to reduce RF inhomogeneity, their approach has the additional, significant advantage of freeing up space inside the bore of the scanner because it does not require close-fitting RF coils. This extra space is extremely valuable because the cost of the large, superconducting magnets needed for MRI increases rapidly with size. Removal of the RF coil thus opens up the possibility of redesigning the scanner's interior to make it more comfortable for patients and/or to reduce cost.

To reap the benefits of the travelling-wave approach, it is important to minimize the reflection and refraction of radio waves that can occur at boundaries between regions of different wave impedance, such as tissue and

air. Brunner *et al.* demonstrate the problems that arise from these effects, but also introduce potential solutions, including the use of wave-impedance matching and absorbers. These involve positioning materials with appropriate electrical properties in the bore of the scanner near to the subject.

The use of waveguides comes with a limitation. These structures can only guide signals whose frequencies are above a certain value, the cut-off frequency, which depends on the diameter of the waveguide (the larger the diameter, the lower the cut-off). This places a lower limit on the size of waveguide that can be used for implementation of the travelling-wave approach. At 7 T, there is a fortuitous conjunction of the required NMR frequency

and the frequency cut-off of a waveguide that easily accommodates the human body. Extension of the travelling-wave approach to more commonly available lower-field scanners will require the development of loading arrangements that can reduce the cut-off frequency to encompass lower NMR frequencies.

One potential disadvantage of using the travelling-wave approach for signal detection is that of increased noise pick-up compared with the conventional approach, because the antenna is sensitive to resistive noise sources, such as tissue, or wave absorbers, positioned anywhere in the far field. The high premium placed on the signal-to-noise ratio in MRI means that the optimal technique may involve using the travelling-wave approach for signal excitation and multiple conventional coils for signal detection<sup>4</sup>.

Although Brunner and colleagues' work represents a unique application of the travelling-wave approach to NMR, there are strong analogies with methods used in electron spin resonance and optics. Transfer of ideas that have already been developed for these more mature areas of application should shape future exploitation of the travelling-wave approach in NMR and MRI, and ensure that its full benefits can be realized.

Paul Glover and Richard Bowtell are in the Sir Peter Mansfield Magnetic Resonance Centre, School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK.

e-mail: richard.bowtell@nottingham.ac.uk

1. Hayes, C. E., Edelstein, W. A., Schenck, J. F., Mueller, O. M. & Eash, M. *J. Magn. Reson.* **63**, 622–628 (1985).
2. Brunner, D. O., De Zanche, N., Fröhlich, J., Paska, J. & Pruessmann, K. P. *Nature* **457**, 994–998 (2009).
3. Katscher, U., Börnert, P., Leussler, C. & van den Brink, J. S. *Magn. Reson. Med.* **49**, 144–150 (2003).
4. Pruessmann, K. P., Weiger, M., Scheidegger, M. B. & Boesinger, P. *Magn. Reson. Med.* **42**, 952–962 (1999).

## TAXONOMY

# Three into one will go

Rory Howlett

**Striking instances of larval metamorphosis, and of adult sexual dimorphism, are not uncommon in the animal world. But especially dramatic examples of these phenomena have emerged from the deep sea.**

One corner of the world of ichthyology has just become a little tidier. Writing in *Biology Letters*, Johnson *et al.*<sup>1</sup> report their investigations of representatives of three families of marine fish — which, they find, are in fact the larvae, males and females of a single family. The previous tripartite designation was understandable, however. Johnson and colleagues reveal that the larvae undergo extraordinary metamorphoses, through post-larval and juvenile stages, before becoming transformed into male and female adults that are themselves wondrously different in appearance and anatomy.

The fish concerned are the tapetails, bignose fishes and whalefishes, which respectively are currently assigned to the families Mirapinnidae, Megalomycteridae and Cetomimidae. Johnson *et al.* have explored the possibility that the mirapinnids and megalomycterids are in fact pre-juvenile and adult male cetomimids, respectively. Their investigations involved study of fresh megalomycterids netted at depth in the Gulf of Mexico, as well as of museum holotypes of various forms, and molecular analyses of mitochondrial DNA (a holotype, or type specimen, is the original specimen from which the description of a new taxonomic group is made).

It turns out that the only known and holotype specimen of the deep-sea megalomycterid *Megalomycterus teevani* is actually a ‘transforming mirapinnid’ — that is, a larva caught in the act of metamorphosis into the next stage. Its gonads are histologically male. Similarly, the holotype of the deep-sea mirapinnid *Parataeniophorus gulosus* is an early transforming male specimen, which would also have metamorphosed into a megalomycterid. It therefore seems that the previously ‘missing’ larvae of megalomycterids are mirapinnids.

Could it be that the larvae of cetomimids are also mirapinnids? It seems that the answer is yes. A whalefish, *Cetostoma regani*, caught in the deep southeastern Atlantic, is shown to be a late transforming female morphologically linked with larvae/post-larvae of the mirapinnid *P. gulosus* and males of the megalomycterid *Cetomimoides parri*. This implies that *P. gulosus*

(larvae), *C. regani* (adult females) and *C. parri* (adult males) actually belong to a single taxon, thereby uniting the cetomimids with the two other families.

Johnson *et al.* conclude, then, that tapetails, bignoses and whalefishes are in fact larvae, males and females of members of the family Cetomimidae. This conclusion is consistent with the authors’ evolutionary analysis of mitochondrial-genome sequence data. A task now — and no easy one — will be to find, identify and match up the larvae and adult males with the 20 described species of whalefish. Right now, there do not seem to be enough tapetails (mirapinnids) and bignoses (megalomycterids) to go round.

The extraordinary morphological transformations involved give rise to extreme sexual

dimorphism in the adults. The differences in morphology and habit between larvae, males and females are associated with different feeding mechanisms — or lack of them.

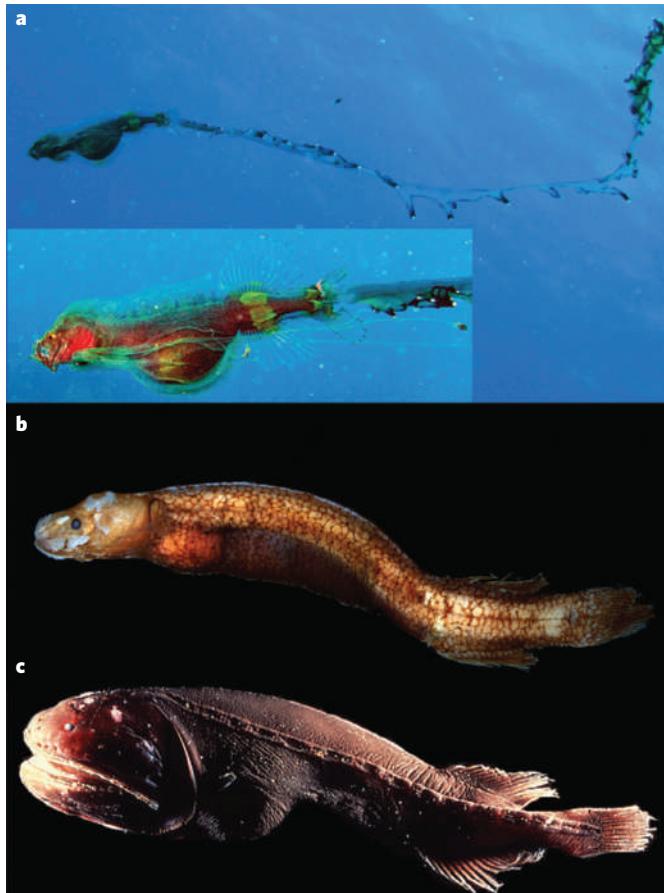
Larvae have small, upturned mouths and feed on copepods in the food-rich, sunlit surface waters. Johnson *et al.*<sup>1</sup> suggest that the long streamers arising from their caudal fin (Fig. 1a) may be involved in feeding or avoiding predators. Males (Fig. 1b) live down in the nutrient-poor bathypelagic zone (depths of 1,000–4,000 metres) and lose the ability to feed: their upper jaws become immobilized, and their stomach and oesophagus disappear. Instead, a ball of copepods in the gut of a transforming male is converted into a massive liver, which then sustains the male through adulthood. Females also live at depth. But they develop specialized gill arches, large horizontal jaws and huge gapes (Fig. 1c). These adaptations allow mature females to capture larger prey and to attain larger body sizes.

Extreme sexual dimorphism is, for example, also found in certain deep-sea anglerfishes. The females look as ‘normal’ as anglerfishes are capable of looking, and do the angling with a lure attached to their head. But the tiny, atrophied males lack a digestive system and live parasitically on the female, with which they fuse<sup>2</sup>; they are nonetheless male enough to provide sperm for reproduction. An example of dimorphism from the invertebrates is that of ‘zombie worms’ of the polychaete genus *Osedax*<sup>3</sup>. Females feed on the bones of whale carcasses that have fallen to the sea floor. But the males — little more than sacks of sperm — live inside the females, and barely develop beyond the larval stage.

As Johnson *et al.*<sup>1</sup> point out, remarkable transformations occur in other deep-sea inhabitants, such as telescopefishes of the family Giganturidae. But a claim they make may well be justified — that the combination of radical larval metamorphosis with extreme adult sexual dimorphism in the now unified Cetomimidae is unparalleled within the vertebrates.

Rory Howlett is at the National Oceanography Centre, Southampton, European Way, Southampton SO14 3ZH, UK.  
e-mail: r.howlett@noc.soton.ac.uk

1. Johnson, G. D. *et al.* *Biol. Lett.* doi:10.1098/rsbl.2008.0722 (2009).
2. Pietsch, T. W. *et al.* *Nature* **256**, 38–40 (1975).
3. Rouse, G. W., Goffredi, S. K. & Vrijenhoek, R. C. *Science* **305**, 668–671 (2004).



**Figure 1 | All whalefish now — larval, male and female cetomimids.**

**a**, The post-larval stage; length is about 6 centimetres not counting the caudal streamer shown in the upper image. **b**, An adult male (5.8 cm). **c**, A juvenile female (7–9 cm).

## MATERIALS SCIENCE

## Let assemblies bloom

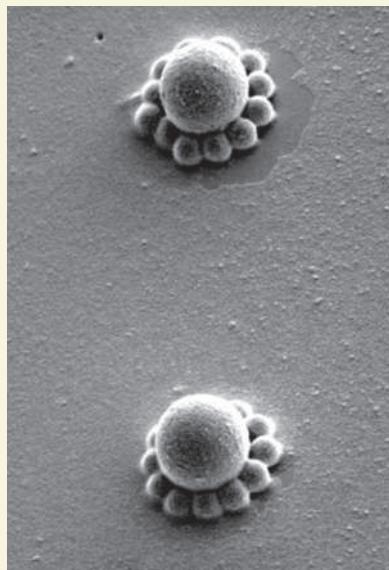
In nature, complexity and order can often result from self-assembly — the spontaneous and reversible aggregation of the components of a disordered system into an ordered one. But so far, researchers have mainly studied systems of one or two components, whereas nature can mix many more — as is the case, for example, with the assembly of ribosomes, the protein-synthesizing organelles in cells. On page 999 of this issue, Erb *et al.* show that man-made multi-component systems can assemble under the action of a magnetic field (R. M. Erb *et al.* *Nature* **457**, 999–1002; 2009).

Self-assembly processes are usually driven by intermolecular forces, including hydrophobicity and the formation of hydrogen bonds, as opposed to the more-stable covalent forces. But external effects such as electrical and magnetic forces can also lend a hand. The components

of such a system can range from metallic nanoparticles to biopolymers and macroscopic objects.

The system studied by Erb and colleagues is a ferrofluid — a mixture of water and magnetic nanoparticles — to which larger, micrometre-sized colloidal particles of different kinds are added. When the fluid is subjected to a magnetic field, the particles can align their magnetic dipoles either with or opposite to the field depending on their magnetic properties with respect to the ferrofluid. When further magnetic nanoparticles are added, what might previously have behaved as a paramagnetic colloidal particle (one that aligns with the field) can become a diamagnetic one (aligning opposite to the field), and vice versa. Such changes cause the particles to assemble into diverse structures.

The type of structure that forms



depends on the size and magnetic properties of the components. Thus, the geometries of the assemblies can feature clusters of smaller particles at the poles of a larger central particle or at its equator (pictured), or even more complex, flower-shaped arrangements.

Erb *et al.* show that these

structures have a remarkably uniform shape, and that they do not aggregate among themselves because their large, central paramagnetic particles repel each other. And they can be made to persist even after the field has been turned off. To do this, their components are made to stick to each other by attaching two kinds of molecule to them that act like Velcro once they are in close proximity. Thus, when the magnetic field is turned off and the structures are dried, they don't fall apart and can be studied individually.

In the future, the use of a larger number of smaller and differently shaped components may help our understanding of complex assemblies to blossom, together with our ability to make them. A possible outcome could be the development of materials with tailored optical properties that would be useful for making biosensors.

**Stefano Tonzani**

R. M. ERB

## MOLECULAR BIOLOGY

## The long and short of RNAs

Piero Carninci

**The known world of RNA is expanding faster than that of any other cellular building block. The latest additions are types of long and short non-coding RNAs formed by bidirectional transcription and unusual processing.**

The relationship between DNA, RNA and protein is no longer as simple as we once thought — that specific genomic sequences are transcribed into messenger RNAs, which are then translated into proteins. In recent years, a flurry of studies has reported the existence of a large and growing family of non-protein-coding RNAs (ncRNAs), describing the different mechanisms of their transcription and their role in regulating gene expression. Six remarkable papers<sup>1–6</sup>, including three in this issue<sup>1–3</sup>, form the latest additions to the list of studies.

Reports of the identification of ncRNAs were initially met with scepticism because of the inherent instability of these single-stranded molecules: RNA is prone to chemical degradation and enzymatic cleavage, as well as to incomplete conversion back to DNA. Sporadic identification of ncRNAs in the past has therefore often been dismissed as one of

these artefacts. High-throughput genomics approaches, however, have begun to chart consistent patterns that are unlikely to be artefacts, and reveal several specific mechanisms for ncRNA biogenesis and processing.

The mechanisms by which ncRNAs are generated seem to be similar to those that operate during mRNA transcription and processing. Normally, the enzyme RNA polymerase II (Pol II) binds to the promoter sequence of a given gene to generate mRNA from the downstream genomic region. In a few cases, the 5' ends of genes on opposite DNA strands are located close together and so share the same promoter, which — when bound to Pol II — bidirectionally triggers their transcription<sup>7,8</sup>. These mRNAs, which are generally long, are also often translated into proteins. The latest studies report the widespread existence of ncRNAs in yeast<sup>2,3</sup> and mammalian cells<sup>1,4,5</sup>, and show that these sequences originate from

known promoters and can also be transcribed in the opposite orientation to known genes.

Promoter-mediated transcription is not the only property that ncRNAs share with mRNA. Soon after transcription begins, a structure called a cap is added to the 5' end of the RNAs transcribed by Pol II (ref. 9); this structure is required for RNA stability and protein translation. Researchers of the Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project<sup>1</sup> (page 1028) show that ncRNAs are also ‘capped’, but — as discussed below — in some circumstances this process seems to occur even independently of transcription.

In the yeast *Saccharomyces cerevisiae*, only a few RNA sequences, called cryptic unstable transcripts (CUTs), had previously been identified. Xu *et al.*<sup>2</sup> (page 1033) and Neil *et al.*<sup>3</sup> (page 1038) now show that these sequences form a main class of Pol-II-transcribed long ncRNAs, which are degraded soon after synthesis by an RNA surveillance pathway. By preventing degradation of CUTs experimentally, the authors show that these ncRNAs are a consistent part of the cell's total pool of transcripts, being transcribed from promoter regions in both directions. In fact, the authors find that at least a third of all genes are transcribed into either CUTs or another, more stable class of long ncRNAs called SUTs (for stable unannotated transcripts) (Fig. 1a). Until now, the only

function ascribed to CUTs has been control of gene expression, for instance by silencing transcription through deacetylation of histone proteins<sup>10</sup>. On the basis of some direct and indirect evidence, the two teams<sup>2,3</sup> propose that CUTs, and indeed SUTs, might be broad regulators of transcription in yeast.

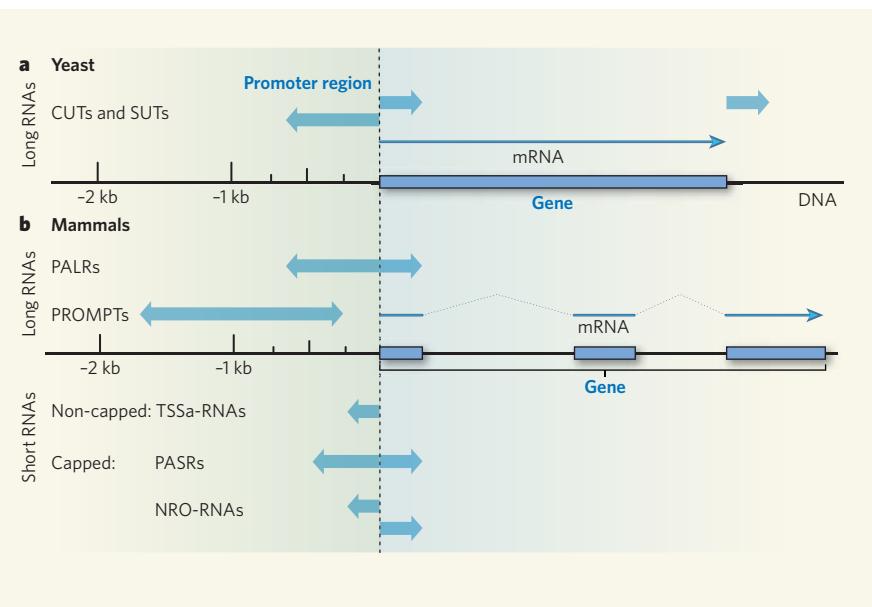
In a parallel study of human cells<sup>6</sup>, a class of ncRNAs called promoter upstream transcripts (PROMPTs) was discovered, which resemble yeast's CUTs. In addition, PROMPTs partially overlap — in size and distance from promoter regions — with a previously identified<sup>11</sup> class of human ncRNAs called promoter-associated long RNAs (PALRs); PROMPTs are, however, less stable than PALRs.

Like mRNAs, ncRNAs are normally transcribed in the downstream direction relative to a promoter. Surprisingly, however, certain PROMPTs are transcribed from upstream sequences towards promoters and their associated genes, but apparently lack their own promoters<sup>6</sup>. The identification of RNA transcription in the absence of promoters is puzzling and raises further questions about ncRNA biogenesis. Are PROMPTs transcribed from unrecognized promoters? Or perhaps they use their antisense RNA sequence as a template: the antisense RNAs could still be generated by bidirectional promoters, which produce PROMPTs antisense to known mRNAs. This possibility, however, would require the action of an RNA-dependent RNA polymerase, which has not been detected in mammals.

The studies also explored the biogenesis of short mammalian ncRNAs<sup>1,4,5</sup>. Members of the ENCODE Transcriptome Project<sup>1</sup> show that constituents of a class of short RNAs that they had identified previously — promoter-associated short RNAs (PASRs) — are capped, as these sequences often overlap with cap markers called CAGE tags; until now, these tags were thought to mark only transcription start sites<sup>12</sup>. Moreover, the authors' further analysis of PASRs and other short RNAs indicates that, as in yeast, mammalian promoters mediating transcription of these sequences are also inherently bidirectional.

Two other teams<sup>4,5</sup> also reached this conclusion, although with different classes of short ncRNAs: TSSa-RNAs and NRO-RNAs. These authors show that TSSa-RNAs and NRO-RNAs are transcribed in the opposite direction to the transcripts of known genes and that they overlap with PASRs and PALRs in genomic localization, but differ from them in stability and length. A potential explanation for why only short ncRNAs were detected in these studies<sup>4,5</sup> could be that the longer RNAs are degraded by the RNA surveillance pathway. So, to detect the longer RNA sequences, either RNA degradation should be inhibited or the longer RNAs should be experimentally capped.

Genomic sequences leading to some PASRs and other short RNAs often seem to overlap with those giving rise to PALRs, as well as with CAGE tags<sup>1</sup>. Nonetheless, some of these short



**Figure 1 | Complexity of transcription<sup>1–6</sup>.** Transcription in eukaryotes is more than just the conversion of genetic code into mRNAs. **a**, In yeast, genomic sequences upstream of gene promoters are transcribed in both directions into long ncRNA sequences called CUTs and SUTs. Single-headed arrows indicate the direction of transcription, and their position indicates the genomic region from which the ncRNA is transcribed. **b**, In the mammalian genome, transcription is even more complex. As well as mRNAs and long ncRNAs such as PALRs and PROMPTs, ncRNAs shorter than 200 nucleotides, including TSSa-RNAs, NRO-RNAs and PASRs, are also transcribed. Some mammalian ncRNAs are capped, which is thought to increase their stability. Double-headed arrows indicate bidirectionality of transcription, with the transcription start site lying anywhere between the arrowheads depending on the specific ncRNA. Arrow length roughly represents the length of the transcripts in the case of long RNAs, and the mapping region for short RNAs.

ncRNAs seem to be further processed by a unique mechanism: they don't seem to acquire their cap during transcription, as do capped coding RNAs. Instead, they are proposed<sup>1</sup> to be generated through a 'secondary capping' process following the cleavage of longer capped RNAs. A subset of the short ncRNAs that are capped in this way belongs to the PASRs, although it is unclear whether some PASRs could also be short primary transcripts that are capped following transcription, without being generated from longer sequences<sup>5</sup>.

Like CUTs and SUTs, the mammalian PASRs might also have regulatory roles. When synthetic PASR sequences corresponding to either the promoter or the sequences within the *MYC* gene were introduced into cells, they inhibited *MYC* transcription, irrespective of the sense or antisense orientation of the PASRs<sup>1</sup>. So PASRs are unlikely to be simply random sequences.

Until the exact molecular mechanisms are clarified, only hypotheses are possible. Transcription bidirectionality is an inherent feature of eukaryotes, and evolution may have found roles for the resulting by-products of gene transcription. Shared promoters might often mediate bidirectional RNA transcription, yet the long antisense RNAs generated in the process might not always be detected because of their instability. Although transcription is generally regulated by the binding of transcriptional activators and repressors to gene promoters<sup>13</sup>, short RNAs — be they capped PASRs or

cleavage products of longer sequences such as PALRs and PROMPTs — might also regulate gene expression by alternative mechanisms.

It is unlikely that every type of ncRNA has now been discovered. Other ncRNA classes — different in size, post-transcriptional modification, relationship with genomic elements, localization in cellular compartments and ultimately function — probably remain to be identified. For starters, a systematic re-examination of the existing data is necessary to classify and unify different names of similar RNA entities. One point is certain, though: the growing role of ncRNA transcription and processing in genome regulation cannot be ignored.

Piero Carninci is at the Omics Science Center, RIKEN Yokohama Institute, Yokohama, Kanagawa 230-0045, Japan.  
e-mail: carninci@riken.jp

- Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project *Nature* **457**, 1028–1032 (2009).
- Xu, Z. *et al.* *Nature* **457**, 1033–1037 (2009).
- Neil, H. *et al.* *Nature* **457**, 1038–1042 (2009).
- Seila, A. C. *et al.* *Science* **322**, 1849–1851 (2008).
- Core, L. J., Waterfall, J. J. & Lis, J. T. *Science* **322**, 1845–1848 (2008).
- Preker, P. *et al.* *Science* **322**, 1851–1854 (2008).
- Trinklein, N. D. *et al.* *Genome Res.* **14**, 62–66 (2004).
- Engstrom, P. G. *et al.* *PLoS Genet.* **2**, e47 (2006).
- Banerjee, A. K. *Microbiol. Rev.* **44**, 175–205 (1980).
- Camblong, J. *et al.* *Cell* **131**, 706–717 (2007).
- Kapranov, P. *et al.* *Science* **316**, 1484–1488 (2007).
- Carninci, P. *et al.* *Nature Genet.* **38**, 626–635 (2006).
- Carninci, P. *et al.* *Curr. Opin. Cell Biol.* **20**, 274–280 (2008).

## OBITUARY

# Frederic Richards (1925–2009)

Pioneer in studies of protein structure and function.

Frederic M. Richards, a leading light in the development of structural biology, died on 11 January at the age of 83.

Richards was a seminal figure in protein science, having had a key role in shaping our fundamental understanding of protein structure and function. His contributions ranged from the early use and development of protein crystallography, to computational analysis of protein geometry, and to the development of chemical probes for the analysis of membrane proteins. Remarkably, throughout his career many of his most influential contributions stemmed from work carried out with his own hands.

Richards came from old New England stock. His undergraduate education at the Massachusetts Institute of Technology was interrupted by service in the army during the Second World War, and was followed by a PhD degree at Harvard Medical School. His first and perhaps most striking contribution began in 1955, during his time as a postdoctoral fellow with Kaj Linderstrøm-Lang at the Carlsberg Laboratory in Copenhagen, and continued after he joined Yale University — where he would remain throughout his scientific career.

Pancreatic ribonuclease A, an enzyme that degrades RNA, was a favoured protein for study in the 1950s. Richards discovered that it could be cut into a short 'S peptide' and a larger 'S protein', which together were termed ribonuclease S and maintained enzyme activity. When these components were separated, each lost all enzymatic activity. But, surprisingly, they could be mixed together to reconstitute the active enzyme, demonstrating that a protein's chemical sequence contains the information needed to attain its active conformation. This work foreshadowed the more extensive refolding studies of ribonuclease A carried out by Nobel laureate Christian Anfinsen. Richards used chemical modification of S peptide to examine the contributions of the amino-acid sequence to the thermodynamic stability of the ribonuclease S complex — all long before the advent of peptide synthesis and site-directed mutagenesis made these types of study widely fashionable.

Together with Harold W. Wyckoff and others, Richards solved the X-ray crystal structure of ribonuclease S and examined the structure with a bound nucleoside monophosphate — this was the second enzyme structure to be solved, and the first protein structure solved in the United States. The relevance of a protein's crystal structure to its conformation in solution

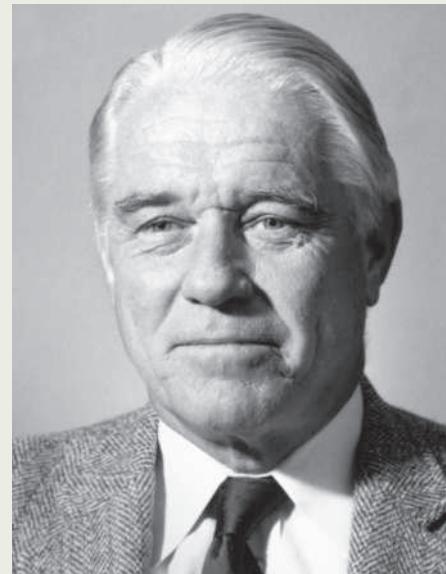
was a serious concern in the early days of protein crystallography. Richards's group established that the ribonuclease S protein was enzymatically active within the crystal, largely putting this issue to rest.

The complex conformations seen in the earliest protein crystal structures were a surprise. Early on, before such crystals were obtained, proteins were believed to be colloidal in nature, and after protein crystals were identified they were expected to display highly symmetrical structures akin to fibrous proteins. Richards developed computational tools to investigate the packing within proteins, leading to the discovery that a protein's core is as well packed as organic molecules in molecular crystals such as table sugar. Richards developed the calculation of the solvent-accessible surface area in proteins and used it to characterize their folding, assembly and function. Numerous investigators have adopted these tools for an ever-expanding variety of applications.

While on sabbatical at the University of Oxford, UK, in 1968, Richards designed and built what became known as the Richards Optical Comparator (or the Richards box), better known in Oxford as Fred's Folly, or simply the Folly, in part because of its architectural similarity to a gazebo. The device contained a half-silvered mirror that allowed the image of a wire model of a protein to be seen floating in the electron-density contour maps determined by X-ray crystallography and drawn on a set of stacked plastic sheets. The Folly permitted the wire model to be manually fitted to the electron density and was quickly adopted by protein crystallographers around the world; it was supplanted only by the advent of molecular-graphics software. My abiding memory of Fred Richards is the grin on his face when I showed him the alamethicin crystals I had grown for my PhD thesis, and his look of pride on seeing the version of the Richards box that I had built to fit the alamethicin model.

The Richards laboratory was a marvellous place for students, postdoctoral fellows and visiting faculty — usually limited to half a dozen, highly independent members pursuing widely disparate projects attacking some fundamental aspect of protein science. Despite the many demands on his time, his office door was always open for energetic and educational scientific discussions. Richards's enthusiasm for protein science was legendary — as were his passions for sailing and ice hockey.

Yale's Department of Molecular Biophysics and Biochemistry was formed by a merger



M. MARSLAND/YALE UNIV.

of existing departments. As founding chair, Richards proved to have a deft touch on this difficult exercise, and showed great insight in hiring an outstanding cadre of junior faculty. The department was remarkable in the breadth of training it required of its graduate students, from molecular biology to molecular biophysics — all long before protein expression allowed both groups to work in the same area.

Richards also played a significant part in many science-policy issues, including his support for the Protein Data Bank (PDB), which was established in 1971 as a site for the voluntary deposition of macromolecular structures and experimental data. Richards led an early steering committee for the PDB and in the mid-1980s was head of an ad hoc committee that successfully encouraged journal editors to require structures reported in their pages to be deposited in the PDB. Richards's committee also succeeded in convincing the US National Institutes of Health to require the deposition of protein coordinates for continued funding. These policy shifts were not without controversy at the time, but their success ultimately facilitated the research of innumerable protein scientists. The timing was crucial, because it put the requirement for deposition in place just before the explosion in the number of structure determinations.

Further insight into Fred Richards's many contributions, and the joy he had making them, can be found in his own words in 'Whatever happened to the fun? An autobiographical investigation', published by the *Annual Review of Biophysics and Biomolecular Structure* in 1997.

## Robert O. Fox

Robert O. Fox is in the Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, Texas 77555-0647, USA.  
e-mail: rofox@utmb.edu

# HYPOTHESIS

## A hierarchical model for evolution of 23S ribosomal RNA

Konstantin Bokov<sup>1</sup> & Sergey V. Steinberg<sup>1</sup>

The emergence of the ribosome constituted a pivotal step in the evolution of life. This event happened nearly four billion years ago, and any traces of early stages of ribosome evolution are generally thought to have completely eroded away. Surprisingly, a detailed analysis of the structure of the modern ribosome reveals a concerted and modular scheme of its early evolution.

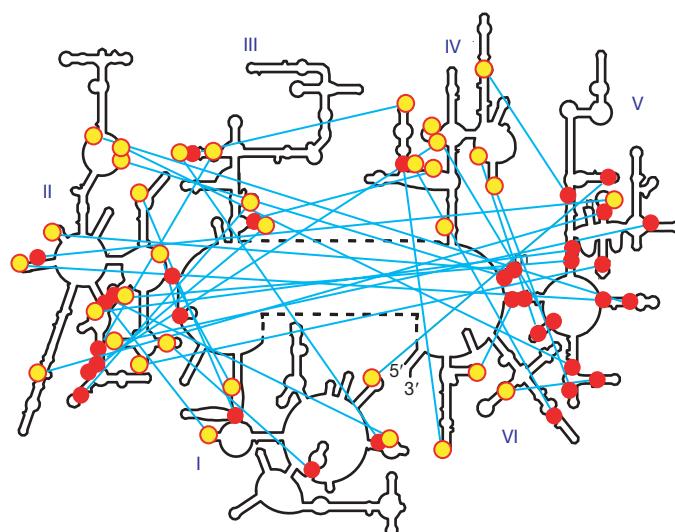
**T**he ribosome is an RNA–protein complex performing protein synthesis in all living cells<sup>1</sup>. It is generally accepted that the ribosome originated from the so-called ‘RNA world’ when proteins did not exist and the primordial chemical reactions of life were catalysed by RNA<sup>2,3</sup>. Although the contemporary ribosome contains several dozen proteins<sup>4–8</sup>, the two major functions of the ribosome—the selection of the proper amino acid and the transpeptidation—are performed by RNA<sup>9–11</sup>, whereas proteins have only an auxiliary role. Structurally, RNA forms the core of the ribosome, whereas proteins are mostly located at the periphery. Hence, the problem of the origin of the ribosome concerns the origin of ribosomal RNA. Because in all living organisms the core of the ribosome has a very similar structure, it must have formed before the split of the tree of life into three phylogenetic domains<sup>12,13</sup>. Consequently, the comparison of the available nucleotide sequences of rRNA is not sufficient for the deduction of how the ribosome emerged. However, the ribosome tertiary structure could provide key clues about the details of this process.

Our initial observation was that, compared to other domains of the 23S rRNA secondary structure, A-minor interactions in domain V<sup>14</sup> follow a very specific pattern. A-minor is a frequently found RNA arrangement consisting of a stack of unpaired nucleotides, predominantly adenosines, that pack with a double helix<sup>15,16</sup>. In the A-minor interactions that domain V forms with other parts of 23S rRNA, the double helix almost exclusively belongs to domain V, whereas the adenosine stack usually belongs to the rest of the molecule (Fig. 1). This characteristic distinguishes domain V from other domains of 23S rRNA, in which the proportion of the adenosine stacks and the double helical regions that form A-minor interactions is reversed.

To explain this abnormality of domain V, we suggest that it reflects the order in which different parts were added to 23S rRNA as it evolved. In the A-minor motif, the conformational integrity of the adenosine stack depends on the presence of the double helix, whereas the helix can maintain a stable conformation without interaction with its counterpart. Presuming that the integrity of the ribosome structure has been maintained throughout its entire evolution, adenosine stacks should not have appeared in rRNA before the corresponding double helices. Because domain V contains the peptidyl-transferase centre (PTC)<sup>17</sup>, which performs the central function of the ribosome, we expect it to be among the most ancient elements of the ribosome structure. Thus, the abnormality of domain V can be explained by the formation of the A-minor interactions between double helices of a more ancient domain V and the adenosine stacks of more recently acquired parts of 23S rRNA.

### The model

The ability of the A-minor motif to serve as an indicator of the relative age of its moieties can be used to determine the order in which different elements were added to the ribosome structure during its evolution. To demonstrate such ordered assembly, we developed a strategy of systematically dismantling the ribosome structure through elimination of those elements that could be considered as most recent acquisitions. As an element, we considered an individual double helix or a domain of stacked nucleotides that on addition to the ribosome structure would form a stable compact arrangement. We suggested a general principle that an element could not be a recent addition if its removal compromised the integrity of the remaining parts of the ribosome. The 5' and 3' ends of a removed element must be structurally close enough to each other to be considered a local insertion. This would guarantee that, after the fragment is removed, the remaining RNA chain maintains its structural integrity. Also, because the integrity of each strand of a double helix



**Figure 1 | Location of inter-domain A-minor interactions in the secondary structure of the *E. coli* 23S rRNA.** The secondary-structure domains are marked by roman numerals. Each A-minor interaction is shown by a cyan line connecting the double helix (red circle) and the corresponding adenosine stack (yellow circle). Unlike other domains, domain V almost exclusively forms these interactions using double helices and not adenosine stacks.

<sup>1</sup>Département de Biochimie, Université de Montréal, Montréal, H3C 3J7 (Québec), Canada.

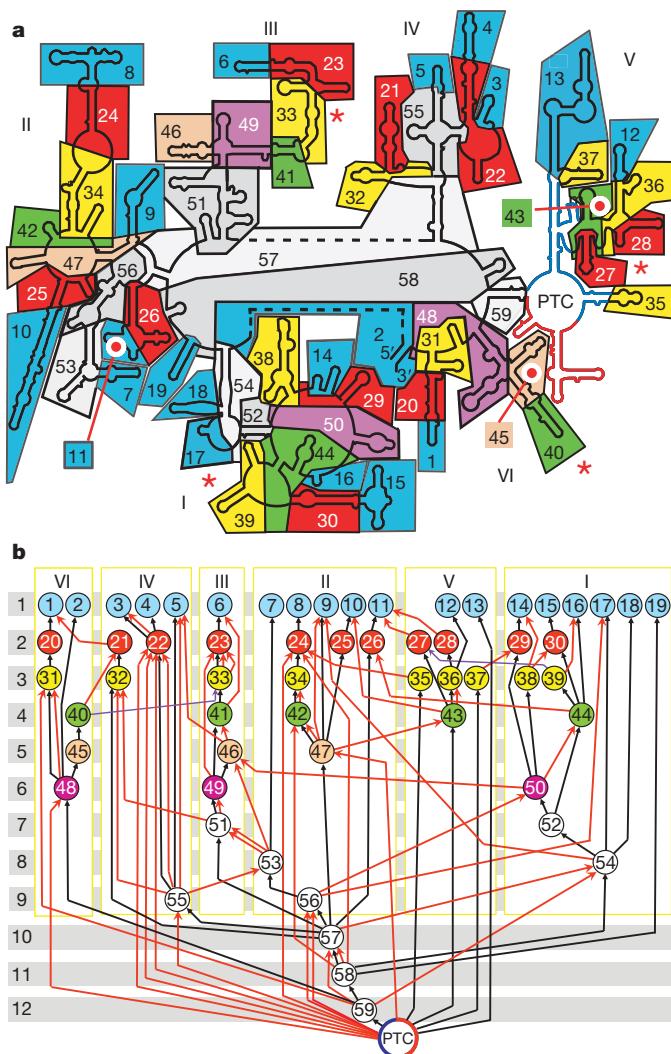
depends on the presence of the other strand, a removed element must contain both strands of the same helix. Finally, if a removed element forms the A-minor motif with the remaining ribosome, it must contain the stack of unpaired nucleotides that form this interaction, and not the double helix. Because our analysis was focused on A-minor interactions, the exact location of the boundaries between different elements was not essential, as long as all adenosine stacks and all corresponding double helical regions remained intact. Additional requirements imposed on elements are discussed in Supplementary Data 1.

Analysis of the tertiary structure of the *Escherichia coli* 23S rRNA<sup>8</sup> revealed 19 elements for which elimination does not compromise the integrity of the remaining part of the structure. These elements form level 1 in Fig. 2b; their location in the 23S rRNA secondary structure is shown in Fig. 2a and their complete description is given in Supplementary Data 1. The identified elements form a total of 13 A-minor interactions with regions located in the remaining part of the molecule (see Supplementary Data 2 and Supplementary Notes 1). In all of these interactions the adenosine stacks belong to the identified elements, whereas the double helices are located in the remaining part of 23S rRNA. Thus, the elements of layer 1 could be considered the final generation of acquired elements. We then identified a further 11 elements, the presence of which is essential for the integrity of only the elements of level 1. Accordingly, we describe these elements as constituting the penultimate generation of added elements (elements of level 2 in Fig. 2b). We repeated the same procedure ten more times and identified a total of 59 elements.

The position and the conformation of each identified element depend on the presence of only the elements of the preceding generations. In Fig. 2b, each dependency of element P on the presence of element Q is shown as arrow Q→P. There are two types of dependencies, D1 and D2. A D1 dependency indicates that the removal of Q before P would split the whole molecule into two separate parts. A D2 dependency indicates that the removal of Q before P would compromise the conformation of P. In total, we identified 59 D1 dependencies and 56 D2 dependencies. Out of all D2 dependencies, 54 were based on the formation of A-minor interactions. The remaining 2 D2 dependencies corresponded to two non-local pseudoknots (discussed later).

The removal of the 12 generations of acquired elements eliminated 93% of the original 23S rRNA. The remaining part, located in domain V, is shown in Fig. 2a by the blue and red lines; its central loop forms the PTC. Recently, it was observed that this region consists of two consecutive parts having practically identical secondary and tertiary structures<sup>11,18</sup> (blue and red parts in Fig. 2a; see also Supplementary Figs 1 and 2). The blue and red parts are arranged symmetrically to each other and form binding sites for the CCA-3'-termini of transfer RNA molecules in the P- and A-sites, respectively. Moreover, there is a very close correspondence between the positions of the nucleotides of both parts involved in the fixation of the equivalent elements of both tRNAs<sup>11,18–21</sup>.

The similarity between both parts is so high that it is logical to suggest that they originated by a duplication of the same RNA fragment. From this point of view, the evolution of 23S rRNA started with an initial fragment of about 110 nucleotides, which, probably, was able to bind the CCA-3'-terminus of what would later be tRNA. The duplication of this fragment allowed the resulting molecule to bind simultaneously two CCA-3'-termini. Within this arrangement, the two CCA-3'-termini associated with both parts are juxtaposed in space to allow for the transpeptidation reaction. Most probably, this dimer was already able to synthesize oligopeptides with random amino acid sequences, which would allow us to call it proto-ribosome. This view is supported by the fact that *in-vitro*-selected small RNA molecules resembling the PTC are able to perform transpeptidation<sup>22</sup>, thus demonstrating that this reaction does not require any other elements of the ribosome structure. All other elements of 23S rRNA were gradually added to the structure, one element at a time, in essentially the same way. Each element could appear only when all elements that were required for its proper positioning had already



**Figure 2 | The location of the identified elements in the *E. coli* 23S rRNA secondary structure (a) and the network of D1 and D2 dependencies between them (b).** Each element has the same colour in a and b. The roman numerals indicate secondary-structure domains. PTC stands for the symmetrical arrangement in domain V containing the peptidyl-transferase centre (the proto-ribosome). a, The two halves of the proto-ribosome are blue and red. Red asterisks indicate the four elements that form two non-local pseudoknots 27–39 and 33–40. b, An arrow connecting two elements Q→P indicates that the position of P depends on the presence of Q. Black and coloured arrows represent D1 and D2 dependencies, respectively. Red arrows Q→P represent A-minor interactions formed by a double helix of element Q and a nucleotide stack of element P. Two violet arrows originate from the dissection of two non-local pseudoknots (see Supplementary Notes 1). The numbers of levels are shown on the left. The detailed description of all elements and of all D2 dependencies is given in Supplementary Data 1 and 2.

been placed. New elements were added as insertions containing all necessary details to dock with the surface of the evolving ribosome without disturbing already existing parts. The most common way for a new element to be fixed on the ribosome surface would be through the formation of an A-minor interaction with an already existing double helix.

### Justification of the model

For justification of the suggested evolutionary model, we analysed those features of the 23S rRNA tertiary structure for which the 59 elements could be consecutively removed without damaging the integrity of the remaining part. Our analysis shows that removal of these elements is possible if and only if the arrows representing D1 and D2 dependencies do not form cyclic structures—that is, cases

where a chain of several consecutive arrows arranged head-to-tail starts and finishes at the same element. A mathematically rigorous proof of this statement and the explanation of why the absence of cycles of dependence is essential for dismantling the 23S rRNA structure are given in Supplementary Notes 2.

The absence of cycles built of D1 dependencies reflects the hierarchical topology of the secondary structure of 23S rRNA in which the removal of remote elements of each domain would not compromise the integrity of the remaining RNA chain. Such topology could be disrupted by pseudoknots, several of which exist in 23S rRNA<sup>12</sup>. However, most pseudoknots are arranged in the same region of the secondary structure and can be removed as a single element. Only two pseudoknots between the loops of elements 27–39 and 33–40 are not local (Fig. 2a). However, in both cases it was possible to split the two strands of the inter-loop double helix on the grounds that the conformational integrity of only one of the two loops requires the presence of the other loop (Fig. 2b and Supplementary Notes 1).

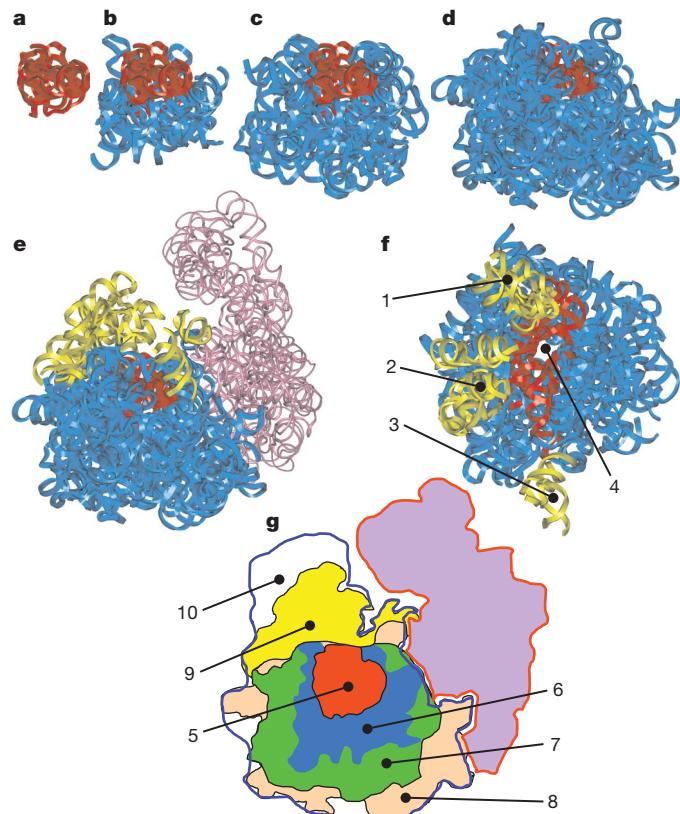
For the given 23S rRNA secondary structure, the absence of cycles involving D2 dependencies is a consequence of the particular orientation of many A-minor interactions. For example, dependency 41→6 (Fig. 2b) stands for the A-minor interaction between the double helix of element 41 and the adenosine stack of element 6. In the opposite situation, if the double helix occurred in element 6, while the adenosine stack was provided by element 41, four elements would have formed cycle 41→33→23→6→41. Similar cycles would have occurred in many other parts of 23S rRNA if the orientations of A-minor interactions were different. The existence of any such cycle would have arrested the procedure of dismantling the 23S rRNA structure before it reached PTC.

To demonstrate how unlikely the absence of cycles really is, we calculated the probability for the 23S rRNA structure to be cycle-free if the orientations of all A-minor interactions were chosen randomly. Our analysis presented in Supplementary Notes 2 shows that the total probability of a cycle-free arrangement in this case would be  $P < 10^{-9}$ . Such low probability excludes the possibility that the absence of cycles of dependence in 23S rRNA has occurred by chance. Instead, it strongly supports a hierarchical scenario for its evolution, according to which the integrity of each element of 23S rRNA depends only on the presence of more ancient elements of its structure. The absence of cycles in the 23S rRNA tertiary structure does not depend on the way we defined individual elements, but instead represents a fundamental property of this molecule.

### Major periods in the 23S rRNA evolution

The scheme of dependencies presented in Fig. 2b can help us to elucidate some details of the evolution of the large ribosomal subunit after the emergence of the proto-ribosome. Our analysis shows that stabilization of the proto-ribosome tertiary structure was a major aspect of the 23S rRNA evolution in the post-proto-ribosome era. In Fig. 3 the structure of the proto-ribosome is shown without other parts of 23S rRNA (Fig. 3a) and with the gradually increasing number of added elements (Fig. 3b–e). The elements forming each structure are shown in Supplementary Fig. 3. The first 8 elements added to the proto-ribosome form a foundation that closely interacts with the bottom part of the proto-ribosome and effectively supports its conformation (Fig. 3b). Further addition of 12 elements makes this foundation wider and more massive (Fig. 3c). Finally, after the addition of a total of 50 elements, the proto-ribosome became surrounded by added elements on all sides except the side from which PTC must be reached by tRNAs (Fig. 3d). The added elements were arranged so they did not interfere with the release of the nascent peptide, leading to the formation of the exit channel (4 in Fig. 3f).

The emergence of the foundation provided new functional opportunities. In particular, it allowed the formation of the area of contact with the small ribosomal subunit (Fig. 3e), which was essential for the integration of this subunit into the ribosome. Another consequence of



**Figure 3 | The aggrandizement of the 23S rRNA structure during its evolution.** **a–e**, the proto-ribosome with 0 (**a**), 8 (**b**), 20 (**c**), 50 (**d**) and all 59 (**e**) elements added. The proto-ribosome is red, elements forming the proto-ribosome foundation are blue, the protuberances are yellow, and 16S rRNA is purple. The complete list of the elements forming structures **a–e** is given in Supplementary Fig. 3. **f**, The top view of the 23S rRNA structure shown in **e**. **g**, The positions of the parts of 23S rRNA shown in **a–e** in the context of the whole ribosome. The structures of the 50S and 30S subunits are contoured by the blue and red line, respectively. 1–3 are the L7/L12, central and L1 protuberances, respectively; 4 is the exit channel; 5–9 are the structures shown in **a–e**, respectively; 10 is the part of 50S subunit that does not include 23S rRNA. This part is formed by ribosomal proteins and 5S rRNA.

the lateral expansion of the proto-ribosome foundation was that it allowed the formation of the three protuberances (yellow in Fig. 3e–g). In Fig. 2b, the elements forming these protuberances are positioned at the upper levels (see Supplementary Fig. 3f). Correspondingly, the particular functions associated with the protuberances—namely, the assistance in the selection of the proper aminoacyl-tRNA and the GTPase reaction<sup>23,24</sup> (the L7/L12 protuberance) as well as in the release of the deacylated tRNA from the E-site<sup>25</sup> (the L1 protuberance)—should be relatively late acquisitions of the ribosome.

Our results also demonstrate that, despite its visible complexity, the structure of 23S rRNA follows a rather simple principle and could have evolved in a relatively short time on the evolutionary scale. Each new insertion emerged randomly and was accommodated only if it made the ribosome more stable and effective as a transpeptidase. At early stages of evolution, the ribosome existed exclusively as an RNA body. Later, when the ribosome functioning became sufficiently effective to produce proteins, the latter started playing an important part in the ribosome structure. We can argue that, among all structures shown in Fig. 3, the structure in Fig. 3b corresponds most closely to the moment when the RNA world changed for the protein-based world. This conclusion is based on the fact that although ribosomal proteins interact with the structure in Fig. 3b only marginally, they form extensive contacts with later structures (not shown). Whether indeed the structure in Fig. 3b corresponds to the end of the RNA world and thus

represents the most effective all-RNA ribosome, however, requires further experimental analysis.

1. Stillman, B. (ed.) *The Ribosome*. Cold Spring Harbor Symposia on Quantitative Biology (Cold Spring Harbor Laboratory Press, 2001).
2. Crick, F. H. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–369 (1968).
3. Gilbert, W. The RNA world. *Nature* **319**, 618 (1986).
4. Yusupov, M. M. et al. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
5. Ban, N. et al. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920 (2000).
6. Harms, J. et al. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**, 679–688 (2001).
7. Selmer, M. et al. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).
8. Schuwirth, B. S. et al. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**, 827–834 (2005).
9. Ogle, J. M. et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902 (2001).
10. Noller, H. F., Hoffarth, V. & Zimniak, L. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416–1419 (1992).
11. Nissen, P. et al. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
12. Gutell, R. R., Larsen, N. & Woese, C. R. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol. Rev.* **58**, 10–26 (1994).
13. Doudna, J. A. & Rath, V. L. Structure and function of the eukaryotic ribosome: the next frontier. *Cell* **109**, 153–156 (2002).
14. Cannone, J. J. et al. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2 (2002).
15. Nissen, P. et al. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl Acad. Sci. USA* **98**, 4899–4903 (2001).
16. Doherty, E. A., Batey, R. T., Masquida, B. & Doudna, J. A. A universal mode of helix packing in RNA. *Nature Struct. Biol.* **8**, 339–343 (2001).
17. Polacek, N. & Mankin, A. S. The ribosomal peptidyl transferase center: structure, function, evolution, inhibition. *Crit. Rev. Biochem. Mol. Biol.* **40**, 285–311 (2005).
18. Agmon, I., Bashan, A., Zarivach, R. & Yonath, A. Symmetry at the active site of the ribosome: structural and functional implications. *Biol. Chem.* **386**, 833–844 (2005).
19. Samaha, R. R., Green, R. & Noller, H. F. A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. *Nature* **377**, 309–314 (1995).
20. Kim, D. F. & Green, R. Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Mol. Cell* **4**, 859–864 (1999).
21. Hansen, J. L., Schmeing, T. M., Moore, P. B. & Steitz, T. A. Structural insights into peptide bond formation. *Proc. Natl Acad. Sci. USA* **99**, 11670–11675 (2002).
22. Zhang, B. & Cech, T. R. Peptide bond formation by *in vitro* selected ribozymes. *Nature* **390**, 96–100 (1997).
23. Savelsbergh, A. et al. Stimulation of the GTPase activity of translation elongation factor G by ribosomal protein L7/12. *J. Biol. Chem.* **275**, 890–894 (2000).
24. Kavran, J. M. & Steitz, T. A. Structure of the base of the L7/L12 stalk of the *Haloarcula marismortui* large ribosomal subunit: analysis of L11 movements. *J. Mol. Biol.* **371**, 1047–1059 (2007).
25. Nikulin, A. et al. Structure of the L1 protuberance in the ribosome. *Nature Struct. Biol.* **10**, 104–108 (2003).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank L. Brakier-Gingras, A. Mankin, S. Michnick and I. Ponomarenko for advice and comments. This work was supported by a grant from NSERC.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.V.S. ([serguei.chteinberg@umontreal.ca](mailto:serguei.chteinberg@umontreal.ca)).

## ARTICLES

# APP binds DR6 to trigger axon pruning and neuron death via distinct caspases

Anatoly Nikolaev<sup>1</sup>, Todd McLaughlin<sup>2</sup>, Dennis D. M. O'Leary<sup>2</sup> & Marc Tessier-Lavigne<sup>1</sup>

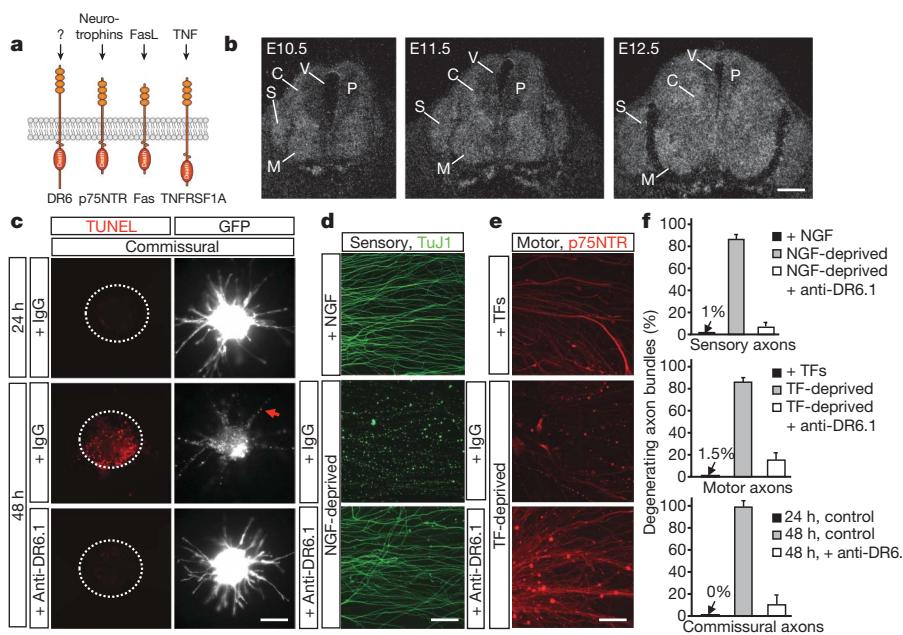
Naturally occurring axonal pruning and neuronal cell death help to sculpt neuronal connections during development, but their mechanistic basis remains poorly understood. Here we report that  $\beta$ -amyloid precursor protein (APP) and death receptor 6 (DR6, also known as TNFRSF21) activate a widespread caspase-dependent self-destruction program. DR6 is broadly expressed by developing neurons, and is required for normal cell body death and axonal pruning both *in vivo* and after trophic-factor deprivation *in vitro*. Unlike neuronal cell body apoptosis, which requires caspase 3, we show that axonal degeneration requires caspase 6, which is activated in a punctate pattern that parallels the pattern of axonal fragmentation. DR6 is activated locally by an inactive surface ligand(s) that is released in an active form after trophic-factor deprivation, and we identify APP as a DR6 ligand. Trophic-factor deprivation triggers the shedding of surface APP in a  $\beta$ -secretase (BACE)-dependent manner. Loss- and gain-of-function studies support a model in which a cleaved amino-terminal fragment of APP (N-APP) binds DR6 and triggers degeneration. Genetic support is provided by a common neuromuscular junction phenotype in mutant mice. Our results indicate that APP and DR6 are components of a neuronal self-destruction pathway, and suggest that an extracellular fragment of APP, acting via DR6 and caspase 6, contributes to Alzheimer's disease.

The initial formative phase of nervous system development, involving the generation of neurons and extension of axons, is followed by a regressive phase in which inappropriate axonal branches are pruned to refine connections, and many neurons are culled to match the numbers of neurons and target cells<sup>1–3</sup>. The loss of neurons and branches also occurs in the adult after injury, and underlies the pathophysiology of many neurodegenerative diseases<sup>1,4</sup>.

Our understanding of regressive events in development remains fragmentary. Degeneration can result 'passively' from the loss of support from trophic factors such as nerve growth factor (NGF)<sup>1–3</sup>.

There is also evidence for 'active' mechanisms in which extrinsic signals trigger degeneration by means of pro-apoptotic receptors, including some members of the tumour necrosis factor (TNF) receptor superfamily such as p75NTR (also known as NGFR), Fas and TNFRSF1A (previously known as TNFR1) (Fig. 1a)<sup>5</sup>. However, the full complement of degeneration triggers remains incompletely understood.

Our understanding of the intracellular mechanisms of neuronal dismantling is also incomplete. It is well documented that developmental neuronal cell body degeneration requires the apoptotic



**Figure 1 | DR6 regulates degeneration of several neuronal classes.** **a**, Diagram of several TNF receptor superfamily members possessing death domains. **b**, DR6 mRNA is expressed by differentiating spinal neurons (including motor (M) and commissural (C)), and by sensory (S) neurons in DRG at E10.5–E12.5. Expression is low in neuronal progenitors (P) in the ventricular zone (V). **c–f**, Anti-DR6.1 ( $50 \mu\text{g ml}^{-1}$ ) reduces degeneration *in vitro*. **c**, Anti-DR6.1 inhibits commissural axon degeneration (visualized with green fluorescent protein (GFP), right) and cell body death (TUNEL labelling, left; dots indicate explants) seen after 48 h in dorsal spinal cord cultures. Red arrow indicates degenerating axon. **d, e**, Effect on degeneration of sensory (**d**) or motor (**e**) axons triggered by trophic deprivation. TFs, trophic factors (BDNF and NT3). Axons were visualized by immunostaining for tubulin (TuJ1; sensory) or p75NTR (motor). **f**, The percentage of degenerating axon bundles in **c–e** (mean and s.e.m.,  $n = 3$  replicates). Scale bars, 110  $\mu\text{m}$  (**b, c**) and 50  $\mu\text{m}$  (**d, e**).

<sup>1</sup>Division of Research, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. <sup>2</sup>Molecular Neurobiology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

effectors BAX and caspase 3 (refs 6–8); pruning of a particular dendrite in *Drosophila* is also caspase-dependent<sup>9,10</sup>. Developmental axonal degeneration similarly has many hallmarks of apoptosis—including blebbing, fragmentation, and phagocytic clearing of debris by neighbouring cells<sup>2,4</sup>. However, it has been argued that axonal degeneration is caspase-independent, because caspase 3 inhibitors block cell body but not axonal degeneration<sup>8</sup> (reflecting higher activation of caspase 3 in cell bodies compared to axons<sup>11</sup>), and because genetic manipulations to inhibit apoptosis did not block axonal degeneration in some models<sup>12,13</sup>. These results indicated the existence of a caspase-independent program of axonal degeneration<sup>1,2,4</sup>, but its molecular nature has remained elusive.

While studying the expression of all TNF receptor superfamily members<sup>14</sup>, we found that DR6—one of eight members possessing a cytoplasmic death domain (Fig. 1a)—is widely expressed by neurons as they differentiate and enter a pro-apoptotic state. DR6 is an orphan receptor<sup>15</sup>. In transfected cells, it triggers cell death in a Jun N-terminal kinase-dependent manner<sup>16</sup>. *In vivo*, it regulates lymphocyte development<sup>17,18</sup>, but its involvement in neural development is unknown.

Here we show that DR6 links passive and active degeneration mechanisms. After trophic deprivation, DR6 triggers neuronal cell body and axon degeneration. Because DR6 signals via BAX and caspase 3 in cell bodies, we revisited the involvement of caspases in axonal degeneration, and found that axonal degeneration indeed requires both BAX and a distinct effector, caspase 6. Our results also indicated that DR6 is activated by a prodegenerative ligand(s) that is surface-tethered but released in an active form after trophic deprivation. In searching for candidate ligands with these properties, we considered APP, a transmembrane protein that undergoes regulated shedding and is causally implicated in Alzheimer's disease<sup>19–22</sup>, because we had previously found it to be highly expressed by developing neurons and especially axons (see later). Because Alzheimer's disease is marked by neuronal and axonal degeneration, we had long wondered whether APP participates in developmental degeneration. We show that an extracellular fragment of APP is indeed a ligand for DR6—as is a fragment of its close relative APLP2—that triggers degeneration of cell bodies via caspase 3 and axons via caspase 6, and we propose that this developmental mechanism is hijacked in Alzheimer's disease.

### DR6 regulates neuronal death

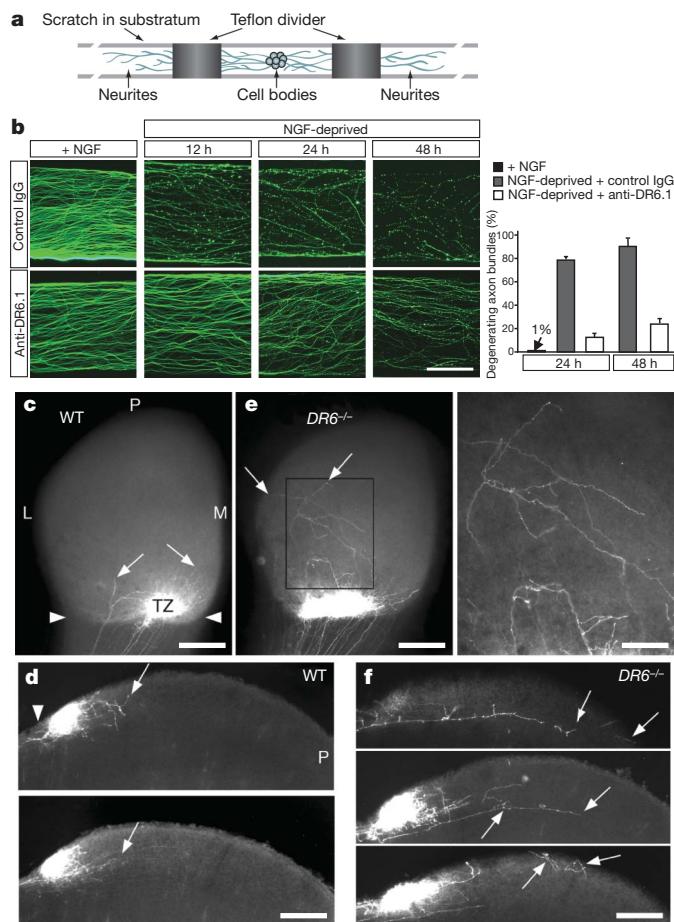
To explore the involvement of the TNF receptor superfamily in neural development, we screened its 28 members by *in situ* hybridization in midgestation mouse embryos. We came to focus on DR6 (Fig. 1a), because its messenger RNA is expressed at low levels in proliferating progenitors in the spinal cord, but is highly expressed by differentiating neurons within the spinal cord and adjacent dorsal root ganglia (DRG) (Fig. 1b).

Because DR6-expressing neurons are becoming dependent for survival on trophic support at these stages, we examined whether DR6 regulates neuronal death after trophic-factor deprivation *in vitro*, focusing on three sets of spinal neurons: commissural, motor and sensory (Supplementary Fig. 1a). Initially, we found that short interfering RNA (siRNA) knockdown of DR6 protected commissural neurons from degeneration (Supplementary Fig. 2). This prompted us to screen monoclonal antibodies to DR6 for their ability to mimic this protection; we selected antibody 3F4 (anti-DR6.1). When embryonic day (E)13.5 rat dorsal spinal cord explants are cultured for 24 h, commissural cell bodies and axons are healthy, but they degenerate if cultured for 24 h longer<sup>23</sup>; anti-DR6.1 inhibited this degeneration (Fig. 1c, f), mimicking DR6 knockdown. Anti-DR6.1 also protected sensory neurons from E12.5 mouse DRGs cultured for 48 h with NGF, and motor neurons from E12.5 mouse ventral spinal cord explants cultured for 24 h with brain-derived neurotrophic factor (BDNF) and neurotrophin 3 (NTF3, also known as NT3): when these cultures were deprived of trophic factor and cultured for 24 h longer, they showed massive cell death and axonal degeneration,

which were largely inhibited by anti-DR6.1 (Fig. 1d–f and Supplementary Fig. 1b). Similar protection was observed when DRGs or ventral explants from a DR6 null mutant<sup>17</sup> were deprived in the absence of anti-DR6.1 (Supplementary Fig. 4b and data not shown), confirming that anti-DR6.1 is function-blocking. DR6 inhibition (by antibody, siRNA or genetic deletion) caused a delay rather than a complete block, because more degeneration was observed in each case 24–48 h later (Fig. 2b, Supplementary Fig. 4b and data not shown). Consistent with a delay, there was a higher motor-neuron number at E14.5 in the DR6 mutant, but this returned to the wild-type level by E18 (Supplementary Fig. 3), after the cell death period. Thus, antagonizing DR6 delays the death of several neuronal populations *in vitro* and *in vivo*.

### DR6 regulates axonal pruning

DR6 protein is expressed not just by cell bodies (data not shown) but also by axons (Supplementary Fig. 4a). Protection of axons by DR6



**Figure 2 | DR6 regulates axon pruning *in vitro* and *in vivo*.** **a**, Diagram of Camponot chamber (adapted from ref. 24). **b**, Images show the local degeneration of sensory axons (TuJ1 immunostain) in Camponot chambers after NGF deprivation from the axonal compartment (top) was delayed by anti-DR6.1 ( $50 \mu\text{g ml}^{-1}$ ) added at the time of deprivation (bottom). The graph shows the percentage of degenerating bundles at 24 and 48 h (mean and s.e.m.,  $n = 3$  replicates). **c–f**, Compromised pruning of retinal axons in *DR6*<sup>-/-</sup> mice. Dorsal view of (**c**, **e**), and vibratome sections through (**d**, **f**), the superior colliculus of wild-type (WT; **c**, **d**) or *DR6*<sup>-/-</sup> (**e**, **f**) mice at P6. In wild-type mice (**c**, **d**), Dil-labelled temporal RGC axons form a dense termination zone (TZ) in anterior superior colliculus (arrowheads denote the anterior border). Few are outside the immediate termination zone area (arrows). In *DR6*<sup>-/-</sup> mice (**e**, **f**), temporal RGC axons and arbor are present in areas far from the termination zone (inset, magnified in **e**, right) and well posterior to it (arrows). L, lateral; M, medial; P, posterior. Scale bars, 100  $\mu\text{m}$  (**b**), 400  $\mu\text{m}$  (**c**, **e**, left), 170  $\mu\text{m}$  (**e**, right) and 250  $\mu\text{m}$  (**d**, **f**).

inhibition might therefore reflect a direct role for DR6 in axons. To explore this, we used compartmented ('Campenot') chambers<sup>24</sup> (Fig. 2a). Sensory neurons are placed in a central chamber containing NGF; their axons grow under a partition into NGF-containing side-chambers. Fluid exchange between the chambers is limited, so NGF deprivation in a side-chamber elicits local axon degeneration while sparing cell bodies<sup>24</sup>. Locally deprived axons degenerate in a stereotypical manner with initial signs by 6 h and extensive degeneration by 12–24 h, but when anti-DR6.1 was added to the deprived side-chamber, degeneration was blocked at 24 h and still largely impaired at 48 h (Fig. 2b). A similar delay was observed when axons of neurons from DR6 knockout mice were locally deprived, but in the absence of anti-DR6.1 (Supplementary Fig. 4b, c). Thus, DR6 function is required in axons for degeneration.

To determine whether DR6 functions in axonal pruning *in vivo*, we studied the well-characterized retino-collicular projection, which develops from an initially exuberant projection of retinal ganglion cell (RGC) axons to a focused termination zone in the superior colliculus. All temporal RGC axons initially extend into posterior superior colliculus, well past their future termination zone in anterior superior colliculus (Supplementary Fig. 5a). This diffuse projection is then refined by selective degeneration of inappropriate axon segments<sup>2</sup>, such that by postnatal day (P)6 in wild-type mice few axon segments persist in areas well beyond the termination zone, as revealed by focal injection of the lipophilic dye DiI into temporal retina (Fig. 2c, d and Supplementary Fig. 5a, b). In contrast, in P6 DR6 mutant mice, many more RGC axons and arbors are present in areas far from the termination zone (Fig. 2e, f and Supplementary Fig. 5c, d): we found an 83% increase in axon-positive domains more than 400 μm from the termination zone (Supplementary Fig. 5f) in DR6<sup>-/-</sup> ( $n=7$ ) compared to wild-type mice ( $n=7$ ;  $P<0.05$ , Student's *t*-test). The defect at P6 represents a delay in pruning, not a complete block, as assessed by examining labelled axons at P4, P5, P6 and P9: at each age, the mutant has more extraneous axons than the wild-type, and fewer are observed in both wild-type and mutant at each age compared to earlier time points, but by P9 the mutant and wild-type projections are indistinguishable (data not shown). Thus, blocking DR6 function delays pruning of sensory axons *in vitro* and retinocollicular axons *in vivo*.

### Caspase 6 regulates axonal degeneration

Because DR6 regulates both cell body apoptosis and axonal degeneration, we revisited whether an apoptotic pathway is also involved in axons. In support, we found that BAX, an effector in the intrinsic apoptotic pathway, is required in axons, because local sensory axon degeneration in Campenot chambers was blocked by the genetic deletion of *Bax* (Fig. 3a) or by local addition of a BAX inhibitor (for example, Supplementary Fig. 10b). Consistent with evidence that caspase 3 mediates cell body but not axon degeneration<sup>8,11</sup>, we found that procaspase 3 is highly enriched in cell bodies, and that zDEVD-fmk, an inhibitor of effector caspases 3 and 7, blocked cell body but not axon degeneration (Fig. 3b, c and Supplementary Fig. 6a–c). There is, however, a third effector caspase, caspase 6. We found that procaspase 6 is expressed in both cell bodies and axons, and that the caspase 6 inhibitor zVEID-fmk blocked degeneration of sensory, motor and commissural axons (Fig. 3b, c and Supplementary Fig. 6a–c), suggesting that caspase 6 regulates axonal degeneration. We verified these results using RNA interference in sensory and commissural neurons: *Casp3* knockdown protected cell bodies significantly but had only a minor protective effect on axons, whereas *Casp6* knockdown protected axons significantly with only minor effect on cell bodies (Fig. 3d, e). Thus, distinct caspases mediate cell body and axon degeneration.

To visualize caspase activation, we first used the fluorescent reporters FAM-DEVD-fmk (for caspase 3 and 7) and FAM-VEID-fmk (for caspase 6), which bind covalently to activated target caspases. In NGF-deprived sensory neurons, the caspase 3/7 reporter labelled cell

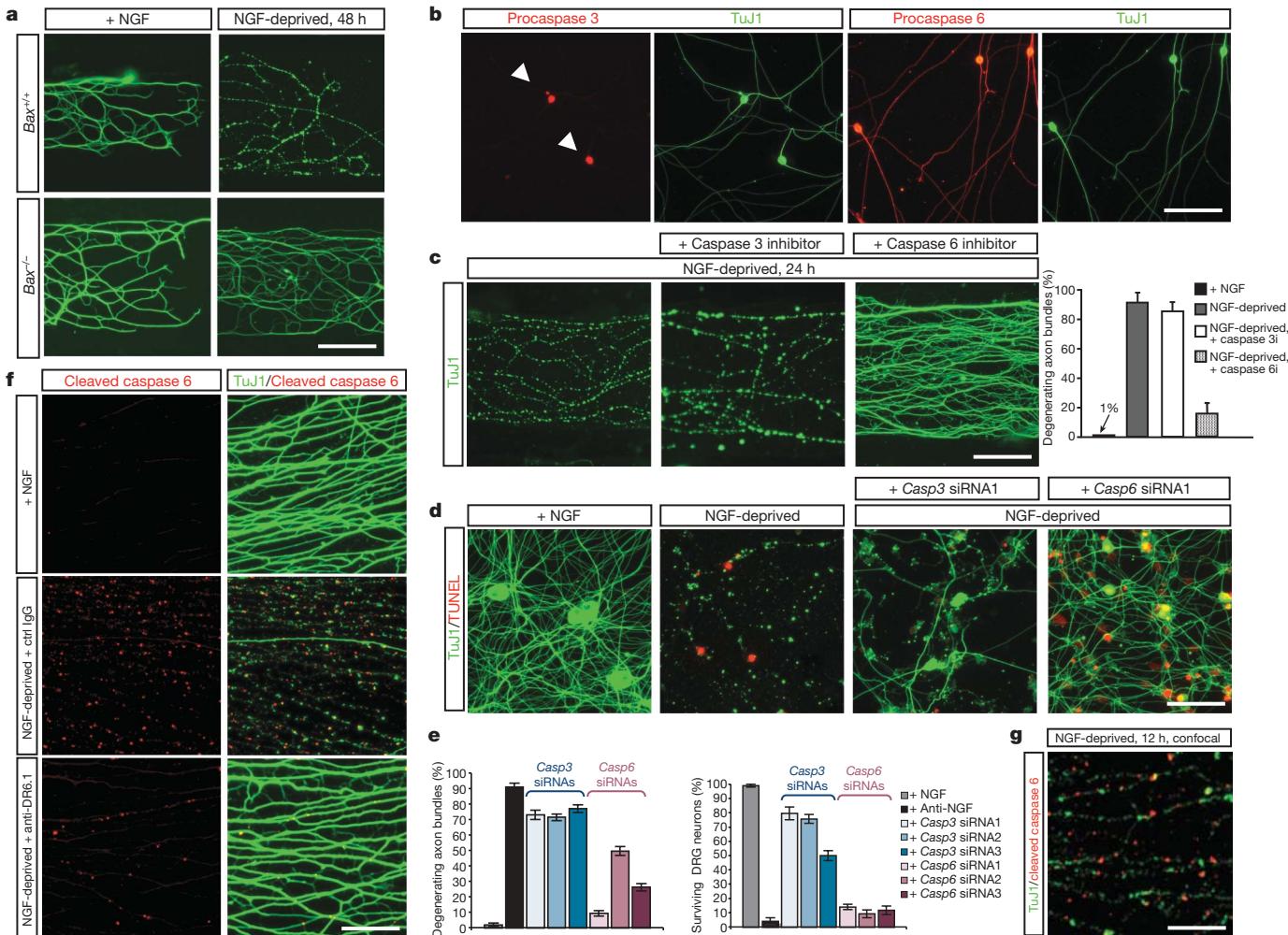
bodies but not axons, consistent with a previous study<sup>11</sup>; in contrast, caspase 6 reporter labelling was observed in both cell bodies and axons, and axonal labelling occurred in regularly spaced 'puncta', giving a beads-on-a-string appearance (Supplementary Fig. 6f). To control for reporter specificity, we used a selective antibody to cleaved caspase 6 and observed a similar punctate pattern in axons (Fig. 3f, g), whereas antibodies to cleaved caspase 3 only label cell bodies<sup>11</sup>. Caspase 6 activation was confirmed biochemically (Supplementary Fig. 6e). Interestingly, caspase 6 activation appeared at sites of microtubule fragmentation (assessed by the loss of tubulin immunoreactivity) (Fig. 3f, g), suggesting that caspase 6 activation drives microtubule destabilization. Punctate caspase 6 activation was markedly reduced by anti-DR6.1 (Fig. 3f) and abolished in *Bax*<sup>-/-</sup> neurons (not shown), suggesting that caspase 6 acts downstream of BAX in the pathway triggered by DR6. However, the possibility of feedback loops in apoptotic pathways makes this interpretation tentative.

### Regulated shedding of a DR6 ligand(s)

As DR6 is a receptor-like protein, we addressed whether it is activated by a ligand(s). If so, the DR6 ectodomain might be capable of binding the ligand(s) and blocking its action (Fig. 4a). Consistent with this, the DR6 ectodomain fused to human Fc (DR6-Fc) mimicked anti-DR6.1 in delaying degeneration (Fig. 4a–c and Supplementary Figs 7a and 13). To search for DR6 binding sites on axons and in conditioned medium, we used the DR6 ectodomain fused to alkaline phosphatase (DR6-AP). Purple alkaline phosphatase reaction product was observed on sensory and motor axons cultured with trophic factors when they were pre-incubated with DR6-AP but not with alkaline phosphatase alone, but binding was markedly reduced after trophic deprivation (Supplementary Fig. 7b, c). To control for the loss of axonal membrane, we blocked degeneration using a BAX inhibitor (data not shown) or using neurons from *Bax*<sup>-/-</sup> mice (Fig. 4d) and observed an even greater reduction in DR6-AP binding (residual binding seen without BAX inhibition might reflect nonspecific binding to degenerating axons). To determine whether DR6-binding sites were shed, we collected medium conditioned by sensory axons (in Campenot chambers) or motor neurons (in explant culture) (a BAX inhibitor was added to prevent nonspecific release resulting from degeneration). Proteins were separated on non-reducing gels, blotted to nitrocellulose, and probed with DR6-AP. Little signal was seen in medium conditioned by either neuronal type in the presence of trophic factor. However, 48 h after trophic deprivation, DR6-AP bound a prominent band around ~35 kDa and a minor band around ~100 kDa in both cultures (Fig. 4e). Together, these results support a 'ligand activation' model in which a prodegenerative DR6 ligand(s) is present on the neuronal surface and inactive, but is shed into medium in an active form after trophic deprivation (Fig. 4f), allowing it to bind and activate DR6.

### N-APP is a regulated DR6 ligand

Several properties of APP made it a candidate for a DR6 ligand: (1) it is highly expressed by developing spinal and sensory neurons and their axons (Fig. 4g), (2) its ectodomain can be shed in a regulated fashion<sup>19,20</sup>, and (3) it is tied to degeneration through its links to Alzheimer's disease<sup>19–22</sup>. In an initial experiment, we found that DR6-AP bound APP expressed in COS-1 cells (Supplementary Fig. 8a). This prompted us to test whether the bands detected by DR6-AP in conditioned medium (Fig. 4e) represent APP ectodomain fragments. APP is cleaved by α- or β-secretases (including, in neurons, BACE1; ref. 25) at distinct sites in its juxtamembrane region (Fig. 4h) to release ~100-kDa ectodomain fragments termed sAPPα or sAPPβ, respectively<sup>19,20</sup>. We probed conditioned medium with a polyclonal antibody to the APP N terminus (anti-N-APP(poly)), which also binds the APP relative APLP2; see later) and an antibody selective for the carboxy-terminal epitope of sAPPβ exposed by BACE cleavage (anti-sAPPβ) (Fig. 4h). Notably, anti-N-APP(poly) detected



**Figure 3 | BAX and caspase 6 regulate axonal degeneration.** **a**, Local sensory axon degeneration (TuJ1 immunostain) 48 h after NGF deprivation in Campenot chambers was blocked in neurons from *Bax*<sup>-/-</sup> mice. **b**, Dissociated sensory neurons double-labelled for pro-caspase 3 and TuJ1 (left), or pro-caspase 6 and TuJ1 (right). Caspase 3 is detected in cell bodies (arrowheads), whereas caspase 6 is seen in both cell bodies and axons. **c**, Local degeneration of sensory axons in Campenot chambers deprived of NGF for 24 h is inhibited by a caspase 6 inhibitor (zVEID-FMK; caspase 6i), but not by a caspase 3/7 inhibitor (zDEVD-FMK; caspase 3i). Quantification is shown to the right (mean ± s.e.m.,  $n = 3$  replicates). **d**, In dissociated sensory neuron cultures deprived of NGF for 24 h, siRNA knockdown of *Casp3* primarily rescues cell body death (TUNEL label), whereas *Casp6* knockdown primarily rescues axonal degeneration. **e**, Quantification of data

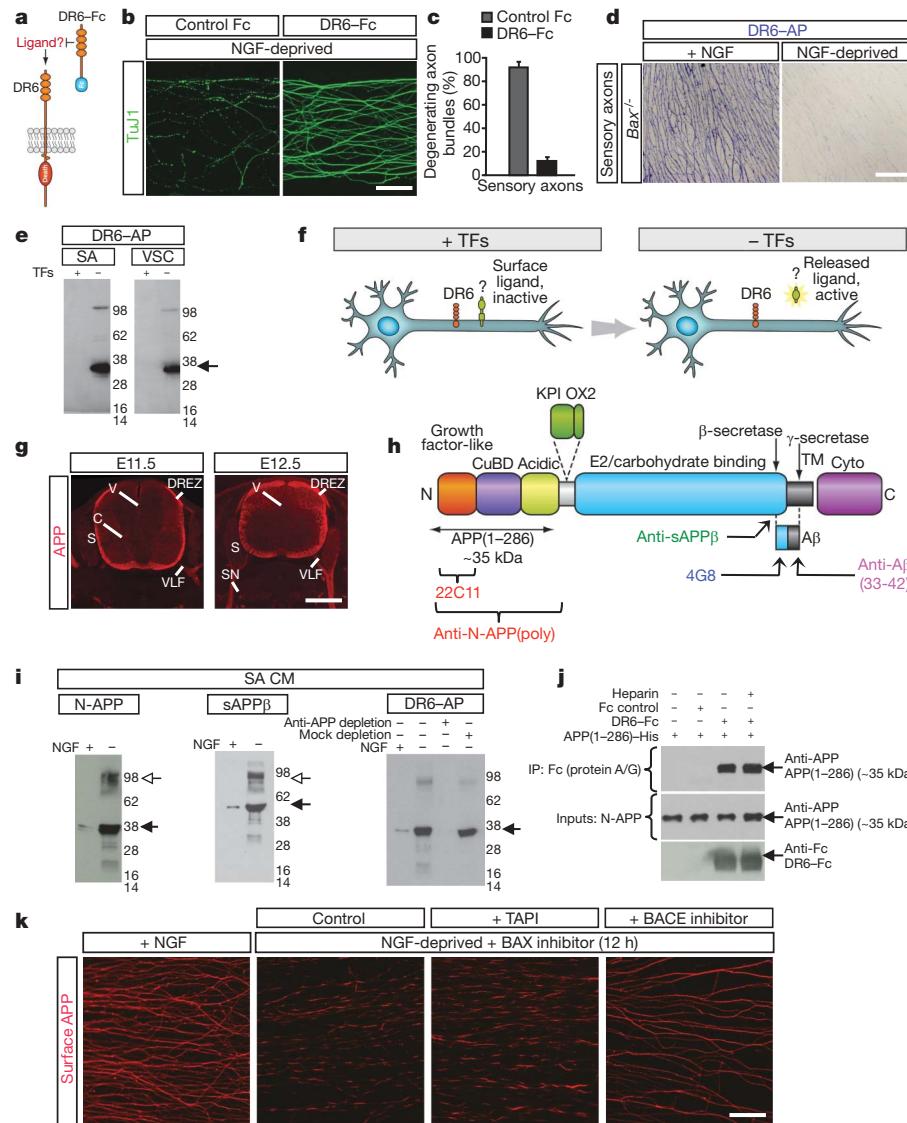
similar bands to those detected by DR6-AP: a major band at ~35 kDa and a minor band at ~100 kDa, both highly enriched after trophic deprivation (Fig. 4i); anti-sAPP $\beta$  detected a minor ~100-kDa band and a major ~55-kDa band (Fig. 4i), also both enriched after trophic deprivation. These results indicate that trophic deprivation triggers BACE cleavage of APP to yield the ~100-kDa sAPP $\beta$  (detected by both antibodies), which undergoes a further cleavage(s) to yield a ~55-kDa C-terminal fragment (detected by anti-sAPP $\beta$ ) and a ~35-kDa N-terminal fragment (detected by anti-N-APP(poly)) that we term N-APP. The site of additional cleavage(s) is unknown, but on the basis of fragment sizes it is expected to be around the junction between the APP ‘acidic’ and ‘E2’ domains (amino acid 286); indeed, recombinant APP(1–286) ran at ~35 kDa and was detected with anti-N-APP(poly) (Fig. 4j), similar to N-APP.

Supporting cleavage of APP by BACE, we found that APP expression on the surface of cultured sensory and motor axons, as assessed with anti-N-APP(poly) and with antibody 4G8 to the APP juxtamembrane region (Fig. 4h), is high in the presence of trophic factor

from **d**. Shown are the percentage of degenerating axon bundles (that is, still visible bundles that show breakdown) and the percentage of surviving neurons (that is, TUNEL $^-$ , TuJ1 $^+$ ) (mean ± s.e.m.,  $n = 3$  replicates). Note that surviving axons may have TUNEL $^+$  cell bodies. The extent of inhibition by individual siRNAs correlates with the degree of target knockdown (Supplementary Fig. 6d). **f**, Detection of caspase 6 activation in sensory axons with a cleaved caspase-6-specific antibody (left; TuJ1 double-label on right). Punctate activation of caspase 6 after NGF deprivation (16 h, middle panel) was reduced by anti-DR6.1 (bottom panel). **g**, Confocal section of a field from **f** shows that active caspase 6 puncta correspond to sites of tubulin loss (fraction non-overlapping:  $82 \pm 3.5\%$ ; mean ± s.e.m.,  $n = 8$  fields). Scale bars, 75  $\mu$ m (**a**, **d**), 100  $\mu$ m (**b**), 50  $\mu$ m (**c**, **f**) and 25  $\mu$ m (**g**).

but lost after trophic deprivation; the surface loss was blocked by three structurally divergent BACE inhibitors—OM99-2, BACE inhibitor IV, and the highly selective AZ29 (ref. 26) but not the  $\alpha$ -secretase inhibitor TAPI (Fig. 4k, Supplementary Figs 9a–c and 10a, and data not shown). Interestingly, 4G8 partially inhibited surface loss (Supplementary Fig. 9d), presumably through the steric hindrance of BACE. Loss of surface APP occurred progressively and in ‘patches’, with little lost at 3 h, more at 6–12 h, and most lost by 24 h (Fig. 4k, Supplementary Fig. 10b and data not shown). Total APP visualized after permeabilization did not change detectably (Supplementary Fig. 10b). Surface loss was not affected by BAX or caspase 6 inhibitors, or in neurons from *Bax*<sup>-/-</sup> mice (Fig. 4k and Supplementary Figs 9c and 10c).

The marked similarity of bands detected by anti-N-APP(poly) and DR6-AP suggested that DR6 binds N-APP. Indeed, depletion of conditioned medium with anti-N-APP(poly) eliminated DR6-AP binding sites (Fig. 4i), and purified DR6-Fc bound to purified recombinant APP(1–286) in pull-down (Fig. 4j) and enzyme-linked



**Figure 4 | The N terminus of APP is a regulated DR6 ligand.** **a**, Diagram of the hypothesis: if DR6 is ligand-activated, then DR6-Fc might sequester ligand and inhibit degeneration. **b**, DR6-Fc inhibits local degeneration of sensory axons in Campenot chambers 24 h after NGF deprivation. **c**, Quantification of results in **b** (mean and s.e.m.,  $n = 3$  replicates). **d, e**, DR6-binding sites are lost from axons and released into medium after trophic deprivation. **f**, DR6-AP binding (purple) to  $Bax^{-/-}$  sensory axons (left) is lost 24 h after NGF deprivation (right). **g**, Medium conditioned by sensory axons (SA) (in Campenot chambers) or ventral spinal cord explants (VSC), maintained with or deprived of trophic factors (TFs) for 24 h (sensory: NGF; motor: BDNF and NT3; BAX inhibitor present), was resolved under non-reducing conditions and probed with DR6-AP. The arrow indicates a major band at  $\sim 35$  kDa. **h**, Results in **a–e** support a ligand activation model in which an inactive DR6 surface ligand is shed in an active form after trophic deprivation. **i**, APP immunostaining on sections of mouse embryos at indicated ages, showing neuronal and axonal expression. DREZ, dorsal root entry zone; S, sensory ganglia; SN, spinal nerve; V, ventricular zone; VLF, ventro-lateral funiculus. **j**, Domain structure of APP (short form, APP695), indicating  $\beta$ - and  $\gamma$ -secretase cleavage sites and antibody binding sites. KPI and OX2 denote alternatively spliced domains of the longer form. Adapted from ref. 20. A $\beta$ , amyloid- $\beta$  peptide; CuBD, copper binding domain; cyto, cytosolic domain; TM, transmembrane domain. **k**, DR6 binding sites in sensory axon conditioned medium (SA CM) include APP ectodomain fragments. Left, anti-N-APP(poly) detects bands at  $\sim 35$  kDa (major; filled arrow) and  $\sim 100$  kDa (minor; open arrow), enriched after trophic deprivation. Middle, anti-sAPP $\beta$  detects bands at  $\sim 55$  kDa (major; filled arrow) and  $\sim 100$  kDa (minor; open arrow). Right, immunodepletion using anti-N-APP(poly) depletes DR6-AP binding sites. Arrow indicates N-APP band at  $\sim 35$  kDa. **l**, Direct interaction between purified APP(1–286) and DR6-Fc revealed by pull-down. The effect of heparin ( $10 \mu\text{g ml}^{-1}$ ) was also examined. IP, immunoprecipitates. **m**, Loss of surface APP in patches from sensory axons 12 h after NGF deprivation is blocked by the BACE inhibitor OM99-2 ( $10 \mu\text{M}$ ) but not by the  $\alpha$ -secretase inhibitor TAPI ( $20 \mu\text{M}$ ). Scale bars, 75  $\mu\text{m}$  (**b, d, k**) and 500  $\mu\text{m}$  (**g**).

ventro-lateral funiculus. **h**, Domain structure of APP (short form, APP695), indicating  $\beta$ - and  $\gamma$ -secretase cleavage sites and antibody binding sites. KPI and OX2 denote alternatively spliced domains of the longer form. Adapted from ref. 20. A $\beta$ , amyloid- $\beta$  peptide; CuBD, copper binding domain; cyto, cytosolic domain; TM, transmembrane domain. **i**, DR6 binding sites in sensory axon conditioned medium (SA CM) include APP ectodomain fragments. Left, anti-N-APP(poly) detects bands at  $\sim 35$  kDa (major; filled arrow) and  $\sim 100$  kDa (minor; open arrow), enriched after trophic deprivation. Middle, anti-sAPP $\beta$  detects bands at  $\sim 55$  kDa (major; filled arrow) and  $\sim 100$  kDa (minor; open arrow). Right, immunodepletion using anti-N-APP(poly) depletes DR6-AP binding sites. Arrow indicates N-APP band at  $\sim 35$  kDa. **j**, Direct interaction between purified APP(1–286) and DR6-Fc revealed by pull-down. The effect of heparin ( $10 \mu\text{g ml}^{-1}$ ) was also examined. IP, immunoprecipitates. **k**, Loss of surface APP in patches from sensory axons 12 h after NGF deprivation is blocked by the BACE inhibitor OM99-2 ( $10 \mu\text{M}$ ) but not by the  $\alpha$ -secretase inhibitor TAPI ( $20 \mu\text{M}$ ). Scale bars, 75  $\mu\text{m}$  (**b, d, k**) and 500  $\mu\text{m}$  (**g**).

immunosorbent assay (ELISA) (Supplementary Fig. 8c) assays. The interaction detected by ELISA is of high affinity (effector concentration for half-maximum response ( $EC_{50}$ ) =  $\sim 4.6$  nM). The interaction of DR6-AP with full-length APP expressed in COS cells was also of high affinity (half maximal saturation =  $\sim 1.3$  nM) (Supplementary Fig. 8a, b). This binding was blocked by anti-N-APP(poly) (data not shown) and anti-DR6.1 (Supplementary Fig. 8a), consistent with APP being a functional DR6 ligand.

Antibodies 4G8 and anti-sAPP $\beta$  used earlier are highly specific for APP. However, like other antibodies to the N terminus of APP<sup>27</sup>,

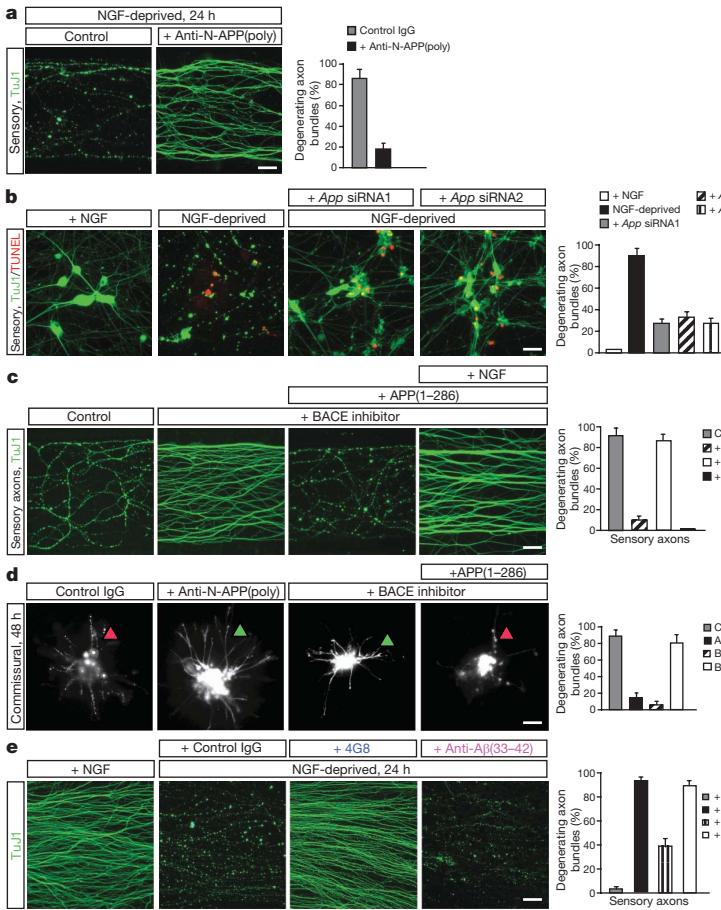
anti-APP(poly) also binds the close APP relative APLP2 (data not shown). We found that a recombinant N-terminal fragment of APLP2 also binds DR6 (Supplementary Fig. 11a). Thus, APLP2 might contribute with APP to the bands detected on western by DR6-AP. Indeed, an antibody selective for the APLP2 N terminus detected a shed fragment in conditioned medium after trophic deprivation (Supplementary Fig. 11b). The relative contribution of APP and APLP2 fragments to DR6-AP binding sites remains to be determined.

To evaluate receptor specificity, we examined by pull-down the binding of APP(1–286) to ectodomains of the seven other

death-domain-containing members of the TNF receptor superfamily, and two orphan members. Only binding to p75NTR was observed (Supplementary Fig. 8d), suggesting that p75NTR might serve as an alternative route for APP effects in some settings; however, the affinity was considerably lower ( $EC_{50} = \sim 300$  nM by ELISA; Supplementary Fig. 8e). Consistent with DR6 being the chief APP receptor in our systems, a fusion of APP(1–286) and alkaline phosphatase bound to sensory axons in culture, but the binding was significantly reduced by anti-DR6.1 or by using *DR6* knockout neurons (Supplementary Fig. 12a, b); residual binding may represent background or binding to another receptor(s), possibly p75NTR.

### Necessity and sufficiency of N-APP

To test whether the N terminus of APP contributes to degeneration, we performed loss-of-function studies. Degeneration of sensory and commissural axons in response to trophic deprivation was inhibited by anti-N-APP(poly) (Fig. 5a, d), which also inhibited the death of sensory neuron cell bodies (Supplementary Fig. 13a, b), without affecting the loss of surface APP after trophic deprivation (Supplementary Fig. 13c). Antibody 22C11 (ref. 28) to the APP N terminus (Fig. 4h) also inhibited sensory axon degeneration (data not shown). Because both antibodies also bind APLP2 (ref. 27), we performed a more selective blockade using RNA interference. Knockdown of *App* in sensory neurons significantly impaired both axon degeneration and cell body death after trophic withdrawal (Fig. 5b). These results support the involvement of an N-terminal fragment of APP in degeneration. In further support, BACE inhibitors impaired degeneration of sensory axons and cell bodies (Fig. 5c) and Supplementary Figs 13a, b and 14) and of commissural axons (Fig. 5d) after trophic deprivation. The selective BACE inhibitor AZ29 blocked degeneration at concentrations consistent with its cellular half-maximal inhibitory concentration ( $IC_{50}$ ) of 470 nM<sup>26</sup> (Supplementary Fig. 14a).



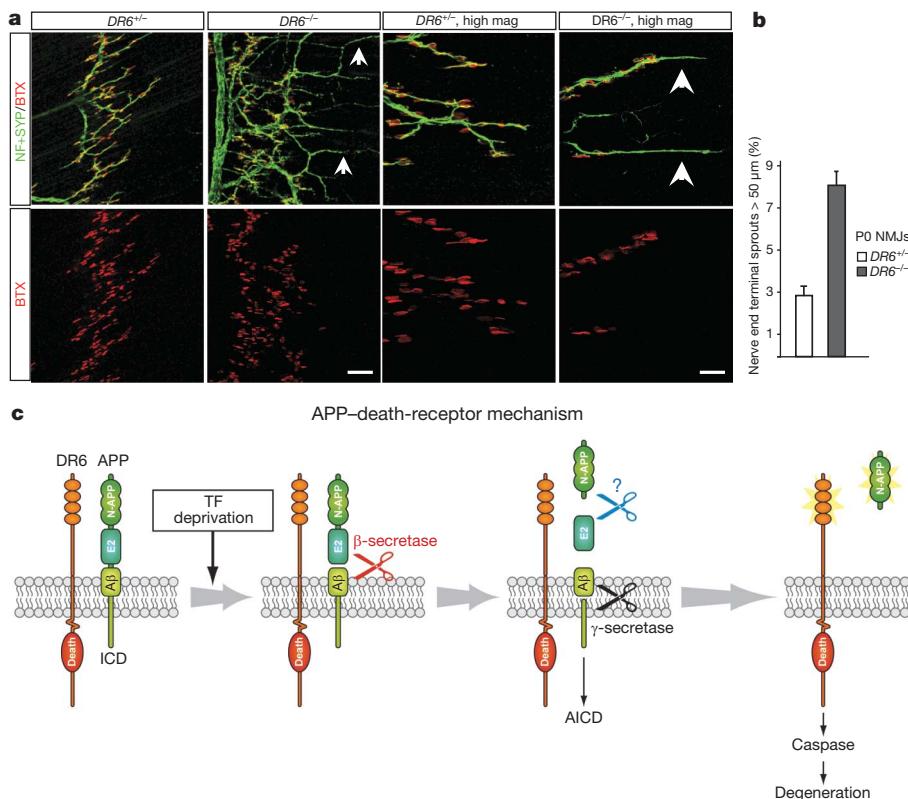
Importantly, axonal degeneration block by BACE inhibitors could be reversed by adding purified APP(1–286) to sensory (Fig. 5c) and commissural (Fig. 5d) neurons, showing that the N terminus of APP is sufficient to trigger degeneration. This effect was largely blocked by anti-DR6.1 (Supplementary Fig. 14b), consistent with DR6 being the most important functional receptor in these cells. Block of sensory cell body degeneration by BACE inhibitors could similarly be reversed by the addition of APP(1–286), albeit at higher concentrations (Supplementary Fig. 13a, b). Together, these results support the model that shed N-APP activates DR6 to trigger degeneration. Degeneration of sensory axons caused by APP(1–286) in the presence of BACE inhibitor was blocked if NGF was present (50 ng ml<sup>-1</sup>; Fig. 5c), indicating that trophic factors also inhibit signalling downstream of DR6.

### Role of amyloid- $\beta$ in physiological degeneration

BACE cleavage of APP is followed by  $\gamma$ -secretase cleavage, yielding amyloid- $\beta$  peptides<sup>19–22</sup> (Figs 4h and 6c). Because amyloid- $\beta$  peptides can be neurotoxic<sup>21,22</sup>, we examined whether they contribute to degeneration. The synthetic amyloid- $\beta$  peptide A $\beta$ (1–42) triggered degeneration in our assays, and an antibody directed at amino acids 33–42 of amyloid- $\beta$  (anti-A $\beta$ (33–42); Fig. 4h) blocked this effect (Supplementary Fig. 9a, e), but did not block degeneration after trophic deprivation (Fig. 5e). Conversely, degeneration induced by synthetic A $\beta$ (1–42) was not blocked by the genetic deletion of *DR6* (data not shown), indicating that amyloid- $\beta$  operates by a mode of action distinct from the physiological degeneration mechanism studied here.

Antibody 4G8 used earlier, which binds amyloid- $\beta$  residues 17–24 (Fig. 4h), also blocked the degenerative effect of A $\beta$ (1–42) (Supplementary Fig. 9e), but unlike anti-A $\beta$ (33–42) it partially inhibited degeneration after trophic deprivation (Fig. 5e). However, as mentioned, 4G8 also partially inhibits the loss of surface

**Figure 5 | The APP N terminus regulates degeneration.** **a**, Local degeneration of sensory axons in Camponot chambers (NGF deprivation, 24 h) was blocked by anti-N-APP(poly) (20  $\mu$ g ml<sup>-1</sup>). Quantification is shown to the right for all panels (**a–e**). **b**, In dissociated sensory neurons, siRNA knockdown of *App* (Supplementary Fig. 13d) significantly reduces axon degeneration 24 h after trophic deprivation, and partially reduces cell body death. **c**, Local degeneration of sensory axons in Camponot chambers (NGF deprivation, 24 h) was inhibited by the local addition of BACE inhibitor (BACEi) OM99-2 (10  $\mu$ M). Purified APP(1–286) added locally restored axonal degeneration, an effect inhibited by 50 ng ml<sup>-1</sup> NGF (right). **d**, Degeneration of commissural neurons and axons at 48 h was inhibited by anti-N-APP(poly) (20  $\mu$ g ml<sup>-1</sup>) or BACE inhibitor OM99-2 (10  $\mu$ M), but restored by APP(1–286). **e**, Effect of amyloid- $\beta$  antibodies on sensory axon degeneration (NGF deprivation, 24 h). 4G8 partially inhibited, whereas anti-A $\beta$ (33–42) did not. Scale bars 50  $\mu$ m (**a**, **c**, **e**), 40  $\mu$ m (**b**) and 200  $\mu$ m (**d**). All data are mean and s.e.m. for  $n = 3$  replicates.



**Figure 6 | APP and DR6 signalling: *in vivo* evidence, and model.** **a**, In control ( $DR6^{+/+}$ ) P0 diaphragm muscle, few axons (green, neurofilament (NF) and synaptophysin (SYP) stain) overshoot endplates (red, fluorescent  $\alpha$ -bungarotoxin (BTX) stain), and those that do are short, but in  $DR6$  mutants more overshoot and many are long (arrowheads). Scale bar, 60  $\mu$ m (left four panels) and 15  $\mu$ m (right four panels). **b**, The number of axons overshooting by  $>50$   $\mu$ m (mean and s.e.m.,  $n = 3$  wild-type, 4 mutants); this underestimates the effect, because overshooting axons are longer in mutants. NMJ, neuromuscular junction. **c**, The APP-death-receptor mechanism is shown. Trophic factor (TF)-deprivation triggers the cleavage of surface APP by  $\beta$ -secretase, releasing sAPP $\beta$ , which is further cleaved by an unknown mechanism (?) to release N-APP. This then binds DR6 to trigger degeneration through caspase 6 in axons and caspase 3 in cell bodies. Also illustrated is cleavage by  $\gamma$ -secretase to release amyloid- $\beta$  (A $\beta$ ) and the APP intracellular domain (AICD).

APP. In contrast, the APP epitope bound by anti-A $\beta$ (33–42) is buried in the cell membrane, so anti-A $\beta$ (33–42) does not bind intact APP nor inhibit its surface loss (Supplementary Fig. 9a, b, d). Because anti-A $\beta$ (33–42) does not protect, we attribute the partial protective effect of 4G8 to its ability to inhibit APP shedding, not its ability to block amyloid- $\beta$  toxicity. Because 4G8 does not bind APLP2, its ability to protect also supports the sufficiency of APP in mediating degeneration.

#### Evidence for an APP-DR6 interaction *in vivo*

To seek evidence for an APP-DR6 interaction *in vivo*, we examined whether the  $DR6$  knockout exhibits any similar phenotype to those reported in the  $App$  knockout—or, given the potential for redundancy, in compound mutants of  $App$  with  $Aplp2$ . One phenotype observed at the neuromuscular junction in the  $App^{-/-} Aplp2^{-/-}$  double knockout is suggestive of a potential pruning defect. In wild-type animals, motor axons normally terminate at synaptic sites (Fig. 6a). In  $App^{-/-} Aplp2^{-/-}$  double knockouts, however, there is a highly penetrant presence of nerve terminals past endplates<sup>29</sup>. Notably, a similar phenotype was observed in the  $DR6$  mutant: rather than terminating at endplates, many terminals were present beyond, giving characteristic finger-like protrusions (Fig. 6a, b). It is not known whether this phenotype reflects a failure to retract or excessive sprouting. Nevertheless, the similarity of phenotypes supports the view that APP signals via DR6 in regulating axonal behaviour *in vivo*. In this system, APP and APLP2 appear redundant because the axonal phenotype is seen only in  $App^{-/-} Aplp2^{-/-}$  double mutants, not single mutants<sup>29</sup>. Whether they are non-redundant in other systems remains to be determined.

#### Discussion

Our results reveal a mechanism, the ‘APP-death-receptor’ mechanism (Fig. 6c), in which trophic deprivation leads to the cleavage of surface APP by  $\beta$ -secretase, followed by further cleavage of the released fragment by an as yet unidentified mechanism (probably near the junction of APP acidic and E2 domains). This then yields

an N-terminal  $\sim 35$ -kDa fragment (N-APP) which binds DR6, triggering caspase activation and degeneration of both neuronal cell bodies (via caspase 3) and axons (via caspase 6). Whether the second cleavage is required for degeneration remains to be determined. Degeneration induced by added APP(1–286) was blocked when sufficient trophic factor was present, indicating that trophic factor not only prevents initiation of the APP cleavage cascade, but also blocks signalling downstream of DR6, providing a fail-safe mechanism to protect if DR6 is inappropriately activated in an otherwise healthy neuron.

#### DR6: an accelerator of self-destruction

In all settings examined, antagonizing DR6 resulted in a delay, rather than a complete block, in neuronal death and axonal pruning. DR6 is therefore best thought of as an accelerator of degeneration—neurons and axons activate it for swift self-destruction when they become atrophic, but without it they have other, slower, ways of achieving that end, perhaps involving other pro-apoptotic receptors<sup>5</sup> or intrinsic mechanisms. This function contrasts that of the DR6 relative p75NTR (which can mediate degeneration when overexpressed<sup>11</sup> or when activated by a neurotrophin in neurons lacking the cognate TRK (also known as NTRK) receptor<sup>5</sup>). p75NTR is more restricted to specific neuronal classes than DR6, and its genetic deletion provided only modest protection of sensory axons in the first 36 h after trophic deprivation (Supplementary Fig. 15), as reported previously for sympathetic axons<sup>30</sup>. In sympathetic neurons, p75NTR is thought to mediate competition for NGF: cells with high NGF/TRKA signalling upregulate expression of BDNF, which acts via p75NTR to trigger degeneration of neighbouring neurons with less robust NGF/TRKA signalling<sup>30,31</sup>. This mechanism shares with ours the expression of a prodegenerative ligand(s) by the neurons themselves. However, the DR6 ligand APP is activated by trophic-factor deprivation, whereas p75NTR ligand expression is increased by trophic-factor stimulation<sup>31</sup>. Thus, p75NTR ligands are released by ‘strong’ neurons to kill ‘weak’ neurons (a paracrine prodegenerative effect)<sup>31</sup>, whereas APP gets activated within weak neurons to accelerate self-destruction

triggered by trophic deprivation or perhaps other insults (an autocrine prodegenerative effect).

### Caspase 6 mediates axonal degeneration

The intracellular mechanisms of axonal degeneration and their relation to apoptosis have been unclear. Our results indicate that developmental axonal degeneration does involve an apoptotic pathway, but with a non-classical effector, caspase 6. Epistasis analysis supports a linear activation model from DR6 to BAX to caspase 6, but does not exclude the possibility that active caspase 6 might feedback, for example, to accelerate the process; in this context, it is intriguing that the APP cytoplasmic domain is a caspase 6 substrate<sup>32</sup>. Activation of caspase 6 by trophic deprivation occurs in a punctate pattern in axons, leading to a beads-on-a-string appearance, and sites of punctate caspase 6 activation correspond to sites of microtubule fragmentation. Caspase 6 might trigger microtubule destabilization by cleaving microtubule associated proteins such as TAU (also known as MAPT), a documented target of caspase 6 (refs 33, 34); in a recent proteomic analysis, almost half the identified caspase 6 targets were cytoskeleton-associated<sup>35</sup>.

### Ligands, receptors for self-destruction

Although p75NTR also binds APP(1–286), DR6 binds with a much higher affinity, and blocking DR6 function largely blocks both APP(1–286) binding to sensory axons and degeneration triggered by APP(1–286). Thus, DR6 seems to be the major functional APP receptor in these neurons, although p75NTR might contribute in other contexts. Conversely, APP may not be the only DR6 ligand: APLP2, which is coexpressed with APP in many neurons<sup>27</sup>, may also contribute to degeneration, because an N-terminal fragment is shed in response to trophic deprivation, can bind DR6, and can trigger degeneration when added exogenously (Supplementary Fig. 16). Future studies will define the relative contributions of APP and APLP2 in different neuronal populations.

The finding of similar neuromuscular junction phenotypes in *DR6* and *App*<sup>−/−</sup> *Aplp2*<sup>−/−</sup> mutants supports a ligand–receptor interaction, and indicates that APP and APLP2 both contribute in this system. The aberrant axonal extensions seen could reflect an impairment of pruning, or, alternatively, a failure of axons to stop; of note, the APP ectodomain has been implicated in neurite growth inhibition<sup>36</sup>. Previous studies have not reported changes in neuronal cell death *in vivo* in *App* mutants, either singly or in combination with *Aplp1* and/or *Aplp2* mutations<sup>37,38</sup>. However, such studies did not examine spinal cord or sensory ganglia, nor involve time-course analysis to evaluate possible delays in degeneration. *In vitro* analysis of cortical neurons from mutants has given divergent results about their basal survival rates and susceptibility to glutamate excitotoxicity<sup>37–39</sup>, but their response to trophic deprivation has not been reported.

In recent findings paralleling ours, trophic deprivation was found to trigger BACE cleavage of APP in PC12 cell-derived neurons and primary hippocampal neurons, and degeneration was reduced by *App* knockdown (in PC12 cells) and BACE inhibition<sup>40,41</sup>. The prodegenerative function of APP was, however, attributed to amyloid-β, because antibodies to amyloid-β inhibited degeneration<sup>40,41</sup>. We too observed protection by antibody 4G8, but attribute this to the ability of 4G8 to bind full-length APP and inhibit cleavage, because a different anti-amyloid-β antibody that does not bind native APP inhibited neither shedding nor physiological degeneration, but blocked the toxic action of added amyloid-β. Conversely, the toxic effect of amyloid-β was not blocked by DR6 inhibition. It was also found that a γ-secretase inhibitor, which reduced amyloid-β production after BACE cleavage, inhibited degeneration<sup>40,41</sup>. We too found that γ-secretase inhibitors can partially inhibit degeneration of commissural and sensory axons (Supplementary Fig. 17), but γ-secretase has many substrates, and it is possible that the efficient activation of DR6 signalling requires a distinct γ-secretase-dependent process. Thus,

our results argue against the involvement of amyloid-β in initiating DR6-dependent degeneration in the neurons studied here, but this does not exclude its possible involvement in other neurons, or at later times in these neurons to augment the effects of APP–DR6 signalling.

### APP–DR6 signalling and neurodegeneration

APP is expressed in adult brain and upregulated in damaged axons<sup>42</sup>. DR6 is also highly expressed in adult brain (Supplementary Fig. 18). Given our findings, it is reasonable to assume that the APP–death-receptor mechanism might contribute to adult plasticity, or to neurodegeneration after injury or in disease. Interestingly, DR6 is upregulated in injured neurons<sup>43</sup>, raising the question as to whether overexpressed DR6 in neurons can trigger ligand-independent degeneration, as reported for p75NTR<sup>11</sup>.

Given the genetic evidence linking APP and its cleavage to Alzheimer's disease, we propose that signalling of APP via DR6 (and possibly p75NTR) may in particular contribute to the initiation or progression of Alzheimer's disease, either alone or in combination with other proposed APP-dependent mechanisms, such as amyloid-β toxicity<sup>21,22</sup> or effects of the APP intracellular domain<sup>44</sup>. Of note, previous studies showed immunoreactivity for the APP N terminus associated with Alzheimer's plaques<sup>45,46</sup>, DR6 maps to chromosome 6p12.2–21.1, near a putative Alzheimer's disease susceptibility locus<sup>47</sup>, and sites of DR6 mRNA expression in adult brain correlate in an intriguing way with known sites of dysfunction in Alzheimer's: very high in hippocampus, high in cortex, but low in striatum (Supplementary Fig. 18), and high in forebrain cholinergic neurons<sup>48</sup>, for instance. In addition, activated caspase 6, a downstream DR6 effector, is associated with plaques and tangles in Alzheimer's disease, and with mild cognitive impairment<sup>34,49</sup>, consistent with the possible activation of caspase 6 in neuritic processes by the APP–death-receptor mechanism (caspase 6 is also implicated in Huntington's disease<sup>50</sup>). Although these results are compatible with the involvement of APP–DR6 signalling in Alzheimer's, it is less clear how the mechanism fits with genetic evidence implicating altered γ-secretase processing in this disease<sup>19–22</sup>. However, the fact that γ-secretase inhibitors antagonize DR6-dependent degeneration hints at a possible relationship.

Thus, further study is required to determine the full implications of the APP–death-receptor mechanism in development, adult physiology and disease. Nonetheless, our results already tie APP to a new mechanism for neuronal self-destruction in development, and suggest that the APP ectodomain, acting via DR6 and caspase 6, contributes to the pathophysiology of Alzheimer's disease.

### METHODS SUMMARY

Antibodies to the following targets were used: procaspase 3 (1:200, Upstate), active caspase 3 (1:200, R&D), procaspase 6 (1:200, Stratagene), active caspase 6 (1:100, BioVision), TuJ1 (1:500, Covance), p75NTR (Chemicon), 2H3 (1:200, DSHB), Islet1/2 (1:100, Santa Cruz Biotech), N-APP (polyclonal, 1:100, Thermo Fisher Scientific; monoclonal 22C11, Calbiochem), DR6 (R&D), NGF (Abcam and Genentech), BDNF (Calbiochem), NT3 (Genentech), amyloid-β (4G8, Covance), the C-terminal cleavage-specific anti-amyloid-β antibody (anti-Aβ(33–42), Sigma), and the C-terminal cleavage site of sAPPβ (Covance). Monoclonal antibodies to human DR6 ectodomain fused with human Fc (A.N., K. Dodge, V. Dixit and M.T.L., manuscript in preparation) were screened for binding to murine DR6 and block of commissural neuron degeneration. Proteins used were: netrin-1 (R&D), NGF (Roche), BDNF and NT3 (Calbiochem), and control IgG (R&D). Transiently expressed murine DR6 ectodomain (amino acids 1–349) fused to human Fc, His-tagged human APP(1–286), and APP(1–286) and APLP2(1–300) fused to human Fc, were affinity-purified from CHO cell supernatants (similar results were obtained using APP(1–286) and APP(1–306), Novus Biologicals). Inhibitors were used against: caspase 3 (Z-DEVD-FMK, Calbiochem), caspase 6 (Z-VEID-FMK, BD Pharmingen), BACE (OM99-2, Calbiochem; BACE inhibitor IV, Calbiochem; AZ29, Genentech), γ-secretase (DAPT, Calbiochem).

See Methods for details of *in situ* hybridization and immunohistochemistry, explant, dissociated, and Campenot chamber cultures, siRNA transfection,

tracing and quantification of retinotectal projections, alkaline-phosphatase-binding assays and pull-down assays.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 24 May; accepted 31 December 2008.

- Raff, M. C., Whitmore, A. V. & Finn, J. T. Axonal self-destruction and neurodegeneration. *Science* **296**, 868–871 (2002).
- Luo, L. & O’Leary, D. D. Axon retraction and degeneration in development and disease. *Annu. Rev. Neurosci.* **28**, 127–156 (2005).
- Buss, R. R., Sun, W. & Oppenheim, R. W. Adaptive roles of programmed cell death during nervous system development. *Annu. Rev. Neurosci.* **29**, 1–35 (2006).
- Saxena, S. & Caroni, P. Mechanisms of axon degeneration: from development to disease. *Prog. Neurobiol.* **83**, 174–191 (2007).
- Haase, G., Pettmann, B., Raoul, C. & Henderson, C. E. Signaling by death receptors in the nervous system. *Curr. Opin. Neurobiol.* **18**, 284–291 (2008).
- Kuida, K. et al. Decreased apoptosis in the brain and premature lethality in CPP32-deficient mice. *Nature* **384**, 368–372 (1996).
- White, F. A., Keller-Peck, C. R., Knudson, C. M., Korsmeyer, S. J. & Snider, W. D. Widespread elimination of naturally occurring neuronal death in *Bax*-deficient mice. *J. Neurosci.* **18**, 1428–1439 (1998).
- Finn, J. T. et al. Evidence that Wallerian degeneration and localized axon degeneration induced by local neurotrophin deprivation do not involve caspases. *J. Neurosci.* **20**, 1333–1341 (2000).
- Kuo, C. T., Zhu, S., Younger, S., Jan, L. Y. & Jan, Y. N. Identification of E2/E3 ubiquitinating enzymes and caspase activity regulating *Drosophila* sensory neuron dendrite pruning. *Neuron* **51**, 283–290 (2006).
- Williams, D. W., Kondo, S., Krzyzanowska, A., Hiromi, Y. & Truman, J. W. Local caspase activity directs engulfment of dendrites during pruning. *Nature Neurosci.* **9**, 1234–1236 (2006).
- Plachta, N. et al. Identification of a lectin causing the degeneration of neuronal processes using engineered embryonic stem cells. *Nature Neurosci.* **10**, 712–719 (2007).
- Sagot, Y. et al. Bcl-2 overexpression prevents motoneuron cell body loss but not axonal degeneration in a mouse model of a neurodegenerative disease. *J. Neurosci.* **15**, 7727–7733 (1995).
- Watts, R. J., Hooper, E. D. & Luo, L. Axon pruning during *Drosophila* metamorphosis: evidence for local degeneration and requirement of the ubiquitin-proteasome system. *Neuron* **38**, 871–885 (2003).
- Bodmer, J. L., Schneider, P. & Tschopp, J. The molecular architecture of the TNF superfamily. *Trends Biochem. Sci.* **27**, 19–26 (2002).
- Bossen, C. et al. Interactions of tumor necrosis factor (TNF) and TNF receptor family members in the mouse and human. *J. Biol. Chem.* **281**, 13964–13971 (2006).
- Pan, G. et al. Identification and functional characterization of DR6, a novel death domain-containing TNF receptor. *FEBS Lett.* **431**, 351–356 (1998).
- Zhao, H. et al. Impaired c-Jun amino terminal kinase activity and T cell differentiation in death receptor 6-deficient mice. *J. Exp. Med.* **194**, 1441–1448 (2001).
- Schmidt, C. S. et al. Enhanced B cell expansion, survival, and humoral responses by targeting death receptor 6. *J. Exp. Med.* **197**, 51–62 (2003).
- Walsh, D. M. et al. The APP family of proteins: similarities and differences. *Biochem. Soc. Trans.* **35**, 416–420 (2007).
- Reinhard, C., Hebert, S. S. & De Strooper, B. The amyloid- $\beta$  precursor protein: integrating structure with biological function. *EMBO J.* **24**, 3996–4006 (2005).
- Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
- Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer’s amyloid  $\beta$ -peptide. *Nature Rev. Mol. Cell Biol.* **8**, 101–112 (2007).
- Wang, H. & Tessier-Lavigne, M. En passant neurotrophic action of an intermediate axonal target in the developing mammalian CNS. *Nature* **401**, 765–769 (1999).
- Camponot, R. B., Walji, A. H. & Draker, D. D. Effects of sphingosine, staurosporine, and phorbol ester on neurites of rat sympathetic neurons growing in compartmented cultures. *J. Neurosci.* **11**, 1126–1139 (1991).
- Cole, S. L. & Vassar, R. BACE1 structure and function in health and Alzheimer’s disease. *Curr. Alzheimer Res.* **5**, 100–120 (2008).
- Edwards, P. D. et al. Application of fragment-based lead generation to the discovery of novel, cyclic amidine  $\beta$ -secretase inhibitors with nanomolar potency, cellular activity, and high ligand efficiency. *J. Med. Chem.* **50**, 5912–5925 (2007).
- Slunt, H. H. et al. Expression of a ubiquitous, cross-reactive homologue of the mouse  $\beta$ -amyloid precursor protein (APP). *J. Biol. Chem.* **269**, 2637–2644 (1994).
- Hilbich, C., Monning, U., Grund, C., Masters, C. L. & Beyreuther, K. Amyloid-like properties of peptides flanking the epitope of amyloid precursor protein-specific monoclonal antibody 22C11. *J. Biol. Chem.* **268**, 26571–26577 (1993).
- Wang, P. et al. Defective neuromuscular synapses in mice lacking amyloid precursor protein (APP) and APP-like protein 2. *J. Neurosci.* **25**, 1219–1225 (2005).
- Singh, K. K. et al. Developmental axon pruning mediated by BDNF-p75NTR-dependent axon degeneration. *Nature Neurosci.* **11**, 649–658 (2008).
- Deppmann, C. D. et al. A model for neuronal competition during development. *Science* **320**, 369–373 (2008).
- LeBlanc, A., Liu, H., Goodyer, C., Bergeron, C. & Hammond, J. Caspase-6 role in apoptosis of human neurons, amyloidogenesis, and Alzheimer’s disease. *J. Biol. Chem.* **274**, 23426–23436 (1999).
- Horowitz, P. M. et al. Early N-terminal changes and caspase-6 cleavage of tau in Alzheimer’s disease. *J. Neurosci.* **24**, 7895–7902 (2004).
- Guo, H. et al. Active caspase-6 and caspase-6-cleaved tau in neurofil thread, neuritic plaques, and neurofibrillary tangles of Alzheimer’s disease. *Am. J. Pathol.* **165**, 523–531 (2004).
- Klaiman, G., Petzke, T. L., Hammond, J. & Leblanc, A. C. Targets of caspase-6 activity in human neurons and Alzheimer disease. *Mol. Cell. Proteomics* **7**, 1541–1555 (2008).
- Young-Pearse, T. L., Chen, A. C., Chang, R., Marquez, C. & Selkoe, D. J. Secreted APP regulates the function of full-length APP in neurite outgrowth through interaction with integrin  $\beta$ 1. *Neural Develop.* **3**, 15 (2008).
- Heber, S. et al. Mice with combined gene knock-outs reveal essential and partially redundant functions of amyloid precursor protein family members. *J. Neurosci.* **20**, 7951–7963 (2000).
- Perez, R. G., Zheng, H., Van der Ploeg, L. H. & Koo, E. H. The  $\beta$ -amyloid precursor protein of Alzheimer’s disease enhances neuron viability and modulates neuronal polarity. *J. Neurosci.* **17**, 9407–9414 (1997).
- Han, P. et al. Suppression of cyclin-dependent kinase 5 activation by amyloid precursor protein: a novel excitoprotective mechanism involving modulation of Tau phosphorylation. *J. Neurosci.* **25**, 11542–11552 (2005).
- Matrone, C. et al. Activation of the amyloidogenic route by NGF deprivation induces apoptotic death in PC12 cells. *J. Alzheimers Dis.* **13**, 81–96 (2008).
- Matrone, C., Ciotti, M. T., Mercanti, D., Marolda, R. & Calissano, P. NGF and BDNF signaling control amyloidogenic route and  $\text{A}\beta$  production in hippocampal neurons. *Proc. Natl. Acad. Sci. USA* **105**, 13139–13144 (2008).
- Medana, I. M. & Esiri, M. M. Axonal damage: a key predictor of outcome in human CNS diseases. *Brain* **126**, 515–530 (2003).
- Chiang, L. W. et al. An orchestrated gene expression component of neuronal programmed cell death revealed by cDNA array analysis. *Proc. Natl. Acad. Sci. USA* **98**, 2814–2819 (2001).
- Muller, T., Meyer, H. E., Egensperger, R. & Marcus, K. The amyloid precursor protein intracellular domain (AICD) as modulator of gene expression, apoptosis, and cytoskeletal dynamics-relevance for Alzheimer’s disease. *Prog. Neurobiol.* **85**, 393–406 (2008).
- Van Gool, D., De Strooper, B., Van Leuven, F. & Dom, R. Amyloid precursor protein accumulation in Lewy body dementia and Alzheimer’s disease. *Dementia* **6**, 63–68 (1995).
- Palmiter, M. R. et al. Antisera to an amino-terminal peptide detect the amyloid protein precursor of Alzheimer’s disease and recognize senile plaques. *Biochem. Biophys. Res. Commun.* **156**, 432–437 (1988).
- Bertram, L. & Tanzi, R. E. Alzheimer’s disease: one disorder, too many genes? *Hum. Mol. Genet.* **13**, R135–R141 (2004).
- Doyle, J. P. et al. Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* **135**, 749–762 (2008).
- Albrecht, S. et al. Activation of caspase-6 in aging and mild cognitive impairment. *Am. J. Pathol.* **170**, 1200–1209 (2007).
- Graham, R. K. et al. Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant huntingtin. *Cell* **125**, 1179–1191 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank R. Axel, C. Bargmann, B. de Strooper, V. Dixit, C. Henderson, J. Lewcock, R. Scheller, R. Vassar, R. Watts, and members of the M.T.-L. laboratory for helpful discussions and suggestions, and critical reading of the manuscript, and A. Bruce for making the diagrams. We thank P. Hass and members of his laboratory (Genentech) for generation and purification of the DR6 ectodomain and APP(1–286), and W.-C. Liang and Y. Wu (Genentech) for binding experiments with purified proteins. Supported by Genentech (A.N. and M.T.-L.) and National Eye Institute grant R01 EY07025 (D.D.M.O.’L.).

**Author Contributions** A.N. performed most of the experiments, with the exception of the analysis of retinal projections and the experiments listed in the Acknowledgements, and co-wrote the paper. The retinotectal analysis was performed by T.M. and supervised by D.D.M.O.’L. M.T.-L. supervised or co-supervised all experiments, and co-wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests; details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.T.-L. (marctl@gene.com).

## METHODS

**In situ hybridization and immunochemistry.** Radioactively labelled  $^{35}\text{S}$  *in situ* mRNA hybridization was as described<sup>51</sup>, using the mRNA locator kit (Ambion). Radiolabelled mRNA probes to antisense sequences of mouse TNF receptor superfamily member 3' untranslated regions were generated using the MAXIscript kit (Ambion). Immunochemistry was as described on sections<sup>51</sup> or cultured cells<sup>52</sup>. Surface labelling was done without detergent. Double labelling was performed using Zenon Technology (Invitrogen). Fluorescent caspase reporter assays were as recommended (MP Biologicals). TUNEL assays were as recommended (*in situ* cell death detection kit, Roche).

**Quantification of axon degeneration.** To measure the percentage of degenerating axon bundles, the number of still visible bundles that showed breakdown was counted.

**Quantification on sections.** Twenty-micrometre serial cryosections were taken from axially matched cervical (C1–C3) levels of  $\text{DR6}^{-/-}$  embryos and heterozygous littermates. Motor neurons were counted in all sections at E14.5 (large Islet1/2-positive ventral neurons; 4 mutants, 3 controls) or E18 (large H&E-stained ventral neurons; 7 mutants, 5 controls).

**Neuronal cultures.** E13 rat dorsal spinal cord was dissected after the introduction of plasmids and siRNAs by electroporation<sup>53</sup>; the dorsal explant survival assay was as described<sup>23</sup>.  $\text{DR6}$  siRNA1 and siRNA2 (sense) were 5'-CAAU-AGGUCAAGGAAGAUGGCU-3' and 5'-AAUCUGUUGAGUCAUGCUCU-3', respectively. The mismatch sequence complementary to siRNA1 was 5'-GGACTCTGTACAGTCACCTCCCAGATCTGTTATAG-3'. Mouse sensory and motor neuron explants or dissociated cells were cultured on laminin-coated 35-mm tissue culture dishes in culture medium (Neurobasal medium with B27 supplement) with appropriate trophic factor (sensory: NGF, 50  $\mu\text{g ml}^{-1}$ ; motor: BDNF and NT3, 10  $\mu\text{g ml}^{-1}$ ). Trophic deprivation was achieved by removing growth factor and adding appropriate antibodies (sensory: anti-NGF, 50  $\mu\text{g ml}^{-1}$ ; motor: anti-BDNF and anti-NT3, 50  $\mu\text{g ml}^{-1}$ ). The introduction of siRNAs into dissociated sensory neuron cultures was performed as described<sup>54</sup>.

**Campenot chamber assay.** The Campenot chamber assay was carried out as described<sup>24</sup> with minor modifications. In brief, 35-mm tissue culture dishes were coated with poly-D-lysine and laminin and scratched with a pin rake (Tyler Research) to generate tracks, as illustrated in Fig. 2a. A drop of culture medium containing 4  $\text{mg ml}^{-1}$  methylcellulose was placed on the scratched substratum. A teflon divider (Tyler Research) was seated on silicone grease and a dab of silicone grease was placed at the mouth of the centre slot. Dissociated sensory neurons from E12.5 mouse DRGs were suspended in methylcellulose-containing medium, loaded into a disposable sterile syringe fitted with a 22-gauge needle, injected into the centre slot under a dissecting microscope, and allowed to settle overnight. The outer perimeter of the dish (the cell body compartment) and the inner axonal compartments were filled with methylcellulose-containing medium. Within 3–5 days *in vitro*, axons begin to emerge into left and right compartments. To trigger local axonal degeneration, NGF-containing medium from axonal compartments was replaced with neurobasal medium containing anti-NGF. Where indicated, anti-DR6.1 or control IgG were added (50  $\mu\text{g ml}^{-1}$  final concentration) at the time of NGF deprivation. Cultures were fixed at different times after deprivation with 4% paraformaldehyde for 30 min at room temperature and processed for TuJ1 immunofluorescence.

**Tracing of RGC axons.** Injections of DiI into temporal retina and subsequent analyses were performed essentially as previously described<sup>55</sup>. The centre of the termination zone was determined to be the centre of a circumscribed circle. Injection size, termination zone size and the efficiency of axon labelling were not different between wild-type and mutant. The termination zone position was also unchanged (average termination zone centre: 50.3% for controls and 49.8% for mutants ( $P > 0.9$ ), from the medial edge). The retina in mutants appeared morphologically normal, with all retinal layers present in similar proportions to wild-type.

For quantification, the presence of axons was determined in 100- $\mu\text{m}$  vibratome sections and transposed onto the superior colliculus. Sections were photographed and axon presence was recorded in 100- $\mu\text{m}$  segments from the anterior border. Using landmarks such as the termination zone, unique arbors, or the edge of the superior colliculus, these data were transposed from photos of sections to photos of the wholemount superior colliculus in dorsal view, resulting in a grid of 100- $\mu\text{m}$  squares covering each superior colliculus. The termination zones and grids were aligned for the analyses. All analyses were performed blinded to genotype.

**Alkaline phosphatase binding assays.** Alkaline phosphatase fused to the DR6 ectodomain (DR6-AP) and to APP(1–286) (APP-AP) were transiently

expressed in COS-1 cells. Medium was changed after 12 h to OPTI-MEM (Invitrogen), and conditioned medium was collected 36 h later and filtered.

The DR6-AP blot assay on conditioned medium was performed as described<sup>56</sup>. In brief, conditioned medium derived from sensory axons maintained in Campenot chambers or ventral spinal cord explants in explant culture (with or without trophic deprivation) was concentrated tenfold using centriprep centrifugal filters (Millipore), resolved in 4–20% gels under non-reducing conditions, and blotted with DR6-AP in alkaline phosphatase binding buffer.

For *in situ* sensory axon binding assays, wild-type or  $\text{Bax}^{-/-}$  sensory explants were cultured, deprived of NGF for 12 h (with or without BAX inhibitor, as indicated), then washed twice with the alkaline phosphatase binding buffer (HBSS, Gibco, with 0.2% BSA, 0.1%  $\text{NaN}_3$ , 5 mM  $\text{CaCl}_2$ , 1 mM  $\text{MgCl}_2$ , 20 mM HEPES, pH 7.0). The alkaline phosphatase binding assay was carried out by making a 1:1 mixture of binding buffer and conditioned medium containing DR6-AP, APP-AP, or control alkaline phosphatase, applied to DRG explants in 8-well culture slides and incubated for 90 min at room temperature. Explants were rinsed five times with binding buffer, fixed with formaldehyde (3.7% in PBS) for 12 min at room temperature, and rinsed three times with HBS (20 mM HEPES, pH 7.0, 150 mM NaCl). Endogenous alkaline phosphatase activity was blocked by heat inactivation at 65 °C in HBS for 30 min. After rinsing three times in alkaline phosphatase reaction buffer (100 mM Tris, pH 9.5, 100 mM NaCl, 50 mM  $\text{MgCl}_2$ ), bound alkaline phosphatase fusion was visualized by developing colour stain in alkaline phosphatase reaction buffer with 1/50 (by volume) of NBT/BCIP stock solution (Roche) overnight at room temperature.

The *in situ* binding of DR6-AP to COS cells transiently expressing APP695 was performed in the same way, but with heparin (2–10  $\mu\text{g ml}^{-1}$ ) added to reduce nonspecific binding. DR6-AP did not bind to controls (p75NTR or DR6 expressed in COS-1 cells) under these conditions. For quantitative analysis, the amount of DR6-AP protein in medium was quantified as described<sup>57</sup>. In brief, 100  $\mu\text{l}$  of 2× alkaline phosphatase buffer (prepared by adding 100 mg para-nitrophenyl phosphate (Sigma) and 15 ml of 1 M  $\text{MgCl}_2$  to 15 ml 2 M diethanolamine, pH 9.8) was mixed with conditioned medium containing DR6-AP or control alkaline phosphatase. The reaction was developed over 5–15 min, with the absorbance being in the linear range (0.1–1). The volume of reaction was adjusted by adding 800  $\mu\text{l}$  of distilled water, and absorbance was measured at 405 nm. Concentration in nM was calculated according to the formula (for 100 ml):  $C(\text{nM}) = A \times 100 \times (60/\text{developing time})/30$ . Saturation binding analysis was performed as described<sup>57</sup>. Prism 4 software (GraphPad) was used to generate saturation binding curves and to determine half-maximal saturation value.

**APP pull-down assays.** Soluble ectodomains of TNF receptor superfamily members fused to human Fc were expressed in CHO cells and affinity purified. They were incubated at 5  $\mu\text{g ml}^{-1}$  in binding buffer (HBSS, with 0.2% BSA, 0.1%  $\text{NaN}_3$ , 5 mM  $\text{CaCl}_2$ , 1 mM  $\text{MgCl}_2$ , 20 mM HEPES, pH 7.0) with 1  $\mu\text{g ml}^{-1}$  of His-tagged APP(1–286) and protein A/G beads (Santa Cruz Bio) at 4 °C overnight. Beads were washed five times with binding buffer. Bound complexes were eluted from beads with SDS loading buffer, resolved in 4–20% SDS PAGE gels, and blotted for APP (with anti-NAPP(poly)) and for TNF receptor family members (with anti-human Fc).

**Mice.** The following mutant mice were used:  $\text{DR6}$  knockout<sup>17</sup> (gift from V. Dixit),  $\text{Bax}$  knockout<sup>58</sup> (gift from S. Korsmeyer) and  $p75\text{NTR}$  knockout<sup>59</sup> (Jackson laboratory).

51. Sabatier, C. et al. The divergent Robo family protein rig-1/Robo3 is a negative regulator of slit responsiveness required for midline crossing by commissural axons. *Cell* **117**, 157–169 (2004).
52. Atwal, J. K. et al. PirB is a functional receptor for myelin inhibitors of axonal regeneration. *Science* **322**, 967–970 (2008).
53. Chen, Z. et al. Alternative splicing of the Robo3 axon guidance Receptor governs the midline switch. *Neuron* **58**, 325–332 (2008).
54. Higuchi, H., Yamashita, T., Yoshikawa, H. & Tohyama, M. Functional inhibition of the p75 receptor using a small interfering RNA. *Biochem. Biophys. Res. Commun.* **301**, 804–809 (2003).
55. McLaughlin, T., Torborg, C. L., Feller, M. B. & O’Leary, D. D. Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. *Neuron* **40**, 1147–1160 (2003).
56. Pettmann, B. et al. Biological activities of nerve growth factor bound to nitrocellulose paper by western blotting. *J. Neurosci.* **8**, 3624–3632 (1988).
57. Okada, A. et al. Boc is a receptor for sonic hedgehog in the guidance of commissural axons. *Nature* **444**, 369–373 (2006).
58. Knudson, C. M. et al.  $\text{Bax}$ -deficient mice with lymphoid hyperplasia and male germ cell death. *Science* **270**, 96–99 (1995).
59. Lee, K. F. et al. Targeted mutation of the gene encoding the low affinity NGF receptor p75 leads to deficits in the peripheral sensory nervous system. *Cell* **69**, 737–749 (1992).

## LETTERS

# Massive star formation within the Leo 'primordial' ring

David A. Thilker<sup>1</sup>, Jennifer Donovan<sup>2</sup>, David Schiminovich<sup>2</sup>, Luciana Bianchi<sup>1</sup>, Samuel Boissier<sup>3</sup>, Armando Gil de Paz<sup>4</sup>, Barry F. Madore<sup>5</sup>, D. Christopher Martin<sup>6</sup> & Mark Seibert<sup>5</sup>

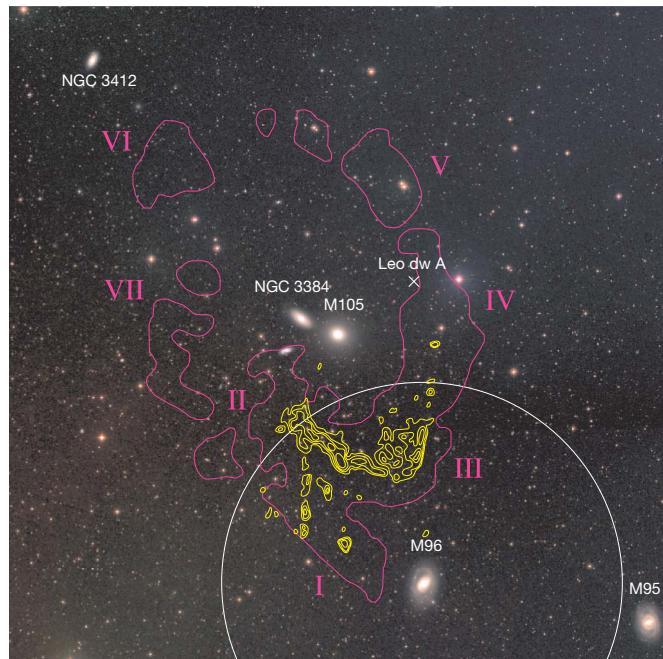
Few intergalactic, plausibly primordial clouds of neutral atomic hydrogen ( $\text{H I}$ ) have been found in the local Universe, suggesting that such structures have either dispersed, become ionized or produced a stellar population on gigayear timescales. The Leo ring<sup>1,2</sup>, a massive ( $M_{\text{HI}} \approx 1.8 \times 10^9 M_{\odot}$ ,  $M_{\odot}$  denoting the solar mass), 200-kpc-wide structure orbiting the galaxies M105 and NGC 3384 with a 4-Gyr period, is a candidate primordial cloud. Despite repeated attempts<sup>3,4</sup>, it has previously been seen only from  $\text{H I}$  emission, suggesting the absence of a stellar population. Here we report the detection of ultraviolet light from gaseous substructures of the Leo ring, which we attribute to recent massive star formation. The ultraviolet colour of the detected complexes is blue, implying the onset of a burst of star formation or continuous star formation of moderate ( $\sim 10^8$ -yr) duration. Measured ultraviolet-visible photometry favours models with low metallicity ( $Z \approx Z_{\odot}/50 - Z_{\odot}/5$ ,  $Z_{\odot}$  denoting the solar metallicity), that is, a low proportion of elements heavier than helium, although spectroscopic confirmation is needed. We speculate that the complexes are dwarf galaxies observed during their formation, but distinguished by their lack of a dark matter component<sup>5</sup>. In this regard, they resemble tidal dwarf galaxies, although without the enrichment preceding tidal stripping. If structures like the Leo ring were common in the early Universe, they may have produced a large, yet undetected, population of faint, metal-poor, halo-lacking dwarf galaxies.

The Leo ring was serendipitously discovered around 25 years ago<sup>1</sup> but remains puzzling today. The most viable interpretation of the ring (Fig. 1) characterizes it as a primordial structure<sup>2,6</sup>, rather than as the product of a gas-sweeping collision<sup>7,8</sup> or any other tidal event involving known group members. Structures like the Leo ring can be formed by a group potential stripping extended gas from a low-surface-brightness disk galaxy<sup>9</sup>, but no such candidate galaxy has been found (the high-surface-brightness spiral galaxy M96 is excluded). Although globally cast as a leftover from group formation, the southern portion of the ring appears affected by interaction with M96, which has drawn some gas out of orbit<sup>10</sup>. Faint  $\text{H}\alpha$  emission was tentatively detected<sup>11</sup> from an ionized component, but this claim was later refuted<sup>12</sup>. Deep broadband optical searches yielded non-detections<sup>3,4</sup>. Such multiwavelength efforts, including molecular gas surveys, are reviewed in ref. 2.

The Leo ring was partially sampled (Fig. 1) by our Galaxy Evolution Explorer (GALEX)<sup>13</sup> Nearby Galaxy Survey observation of M96. A subsequent study will consider the ultraviolet evidence for star formation in other portions of the ring.

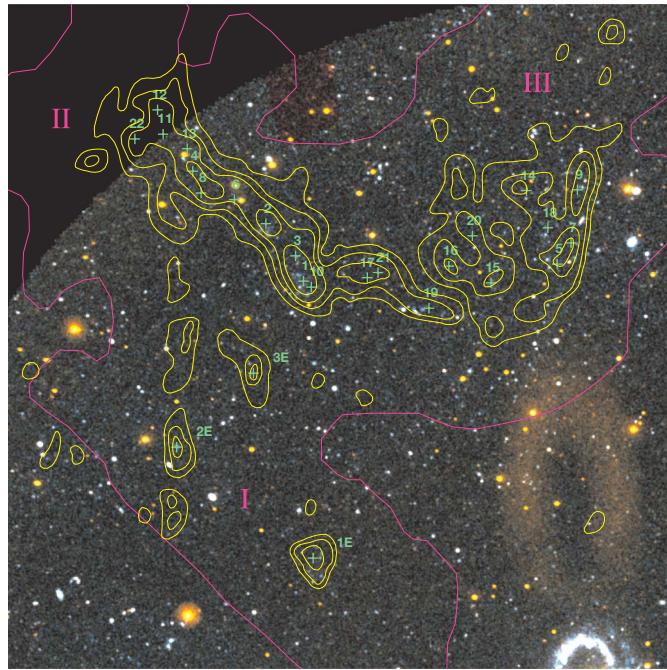
GALEX imaging reveals localized overdensities of sources bright in the far ultraviolet (FUV) within or adjacent to  $\text{H I}$  clumps with neutral

atomic hydrogen column density ( $N_{\text{HI}} > 10^{20} \text{ cm}^{-2}$ ) in the Leo ring. In Fig. 2, we compare the FUV/near-ultraviolet (NUV) image with the  $\text{H I}$  distribution in the southern part of the ring and its extension towards M96. In the ring, we find strong correspondence between ultraviolet and  $\text{H I}$  emission within clump 1 and on the periphery of clump 2 (Fig. 2). We note an enhancement in the surface density of ultraviolet sources, running through clump 18 and offset by  $\sim 1$  arcmin east from clumps 5, 7 and 9. In the M96 extension, GALEX detects a source centred within clump 2E and a slight overdensity at the southern edge



**Figure 1 |  $\text{H I}$  structure of the Leo ring, compared to the distribution of galaxies in the M96 group.** The image is a colour composite spanning two degrees ( $\sim 350$  kpc) and made up from Digitized Sky Survey plates obtained from Skyview (<http://skyview.gsfc.nasa.gov>). The GALEX Nearby Galaxy Survey tile analysed here, NGA\_NGC3368, is marked with a circle. Group members are labelled, including Leo dw A even though its membership is uncertain<sup>10</sup>. The magenta contours ( $N_{\text{HI}} \approx 2 \times 10^{18} \text{ cm}^{-2}$  at 3.3-arcmin resolution) and Roman numerals delineate the large-scale Leo ring  $\text{H I}$  clouds as mapped at Arecibo<sup>2,24</sup>. Very Large Array imaging<sup>5</sup> of the southern ring, showing clumpy substructure, is indicated with yellow contours ( $N_{\text{HI}} = 1.4 \times 10^{19}, 5.5 \times 10^{19}, 1.1 \times 10^{20}, 2.2 \times 10^{20} \text{ cm}^{-2}$  for emission filling the 45-arcsec beam).

<sup>1</sup>Center for Astrophysical Sciences, The Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA. <sup>2</sup>Department of Astronomy, Columbia University, Pupin Physics Laboratories, Mail Code 5246, 550 West 120th Street, New York, New York 10027, USA. <sup>3</sup>Laboratoire d'Astrophysique de Marseille, BP 8, Traverse du Siphon, 13376 Marseille Cedex 12, France. <sup>4</sup>Departamento de Astrofísica, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain. <sup>5</sup>Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, California 91101, USA. <sup>6</sup>California Institute of Technology, MC 405-47, 1200 East California Boulevard, Pasadena, California 91125, USA.

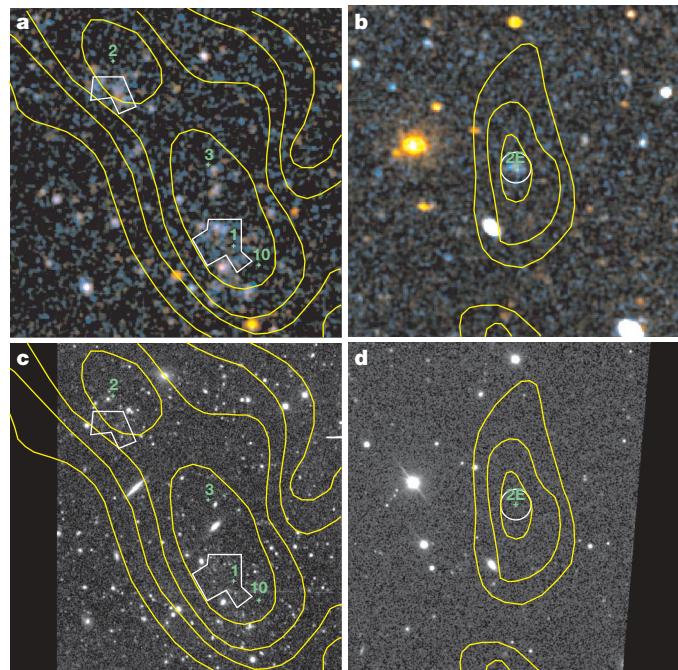


**Figure 2 | Ultraviolet imaging compared with the H I distribution.** Contours are as described for Fig. 1. The field shown spans 35 arcmin ( $\sim 100$  kpc). The Leo ring exhibits considerable H I substructure on arcminute scales<sup>5</sup>, with numerous clumps having H I mass of  $\sim 10^7 M_\odot$  and major axis diameter of a few kiloparsec. Reference 5 tabulated 22 such clumps<sup>5</sup> lying in their Very Large Array survey field of view. These H I clumps (in clouds<sup>24</sup> II and III) are marked with green crosses and labelled. In the M96 extension, ref. 5 did not catalogue any substructures, although several significant peaks were detected in their survey. We designate three structures in the extension which have 45-arcsec resolution, peak  $N_{\text{H}_1}$  exceeding  $10^{20} \text{ cm}^{-2}$  (three contours) as clumps 1E, 2E and 3E. Note the association of ultraviolet emission with H I structures as described in the text. The image is a colour composite of NUV (231.6 nm), energy-weighted average (NUV and FUV) and FUV (153.9 nm) in the R, G, and B colour channels, respectively. The GALEX point spread function full-widths at half-maximum are 4.2 arcsec and 5.3 arcsec in FUV and NUV<sup>13</sup>, respectively, corresponding to  $\sim 200$ – $250$  pc at the assumed 10-Mpc distance<sup>10</sup> to the M96 group. Total exposure time was 1,698 s in both bands. In the ABmag system, the 5 $\sigma$  point-source magnitude limits are 22.6 mag for FUV and 22.7 mag for NUV. This is equivalent to a luminosity of  $1.3 \times 10^{35} \text{ erg s}^{-1}$  at 10 Mpc, which could be produced by a  $10^3 M_\odot$ , 10-Myr-old (or  $2 \times 10^5 M_\odot$ , 100-Myr-old) instantaneous starburst with negligible extinction.

of clump 1E. Imaging of clumps 1 and 2 (Fig. 3) using the 4-m telescope and Mosaic instrument at the Kitt Peak National Observatory (KPNO), Arizona, detects several compact, visible sources ( $r \sim 23$ – $24$  mag) associated with each ultraviolet peak. In Table 1, we provide integrated magnitudes on the scale of the H I clumps.

Some of the  $FUV - NUV$  values (Table 1) are quite blue ( $-0.20$ ,  $-0.31$ ) and imply either young instantaneous bursts ( $10^7$  to a few  $10^7$  yr old) or recent continuous star formation in the age range  $10^8$ – $10^9$  yr. Plausible crossing times (greater than several  $10^7$  yr) for the kiloparsec extent of the star formation complexes eliminate young burst models, on causality grounds.

Figure 4 shows the measured regions on a ( $FUV - NUV$ ,  $FUV - r$ ) colour–colour diagram. Population synthesis predictions<sup>14</sup> for continuous star formation with  $Z = Z_\odot/50$  are plotted using red and black dashed lines depending on the assumed internal extinction. Our current imaging provides limited constraint of extinction. The models become too red in  $FUV - r$  for all Leo ring sources by  $E(B - V) \approx 0.3$  mag, whereas the models are too blue for source 2E if  $E(B - V) = 0.0$  mag. We adopted a representative value of  $E(B - V) = 0.05$  mag, which is slightly less than that determined in extended-ultraviolet-disk star formation regions<sup>15</sup>. The smallest



**Figure 3 | Ultraviolet and visible imaging of detected stellar complexes.** GALEX (a, b; displayed as in Fig. 2) and r-band (c, d) imaging of H I clumps 1 (a, c), 2 (a, c), and 2E (b, d). Both data sets are slightly smoothed and the field shown is 5.43 arcmin ( $\sim 15.8$  kpc). We completed photometry of the ultraviolet-bright regions associated with clumps 1 and 2 using carefully selected polygonal apertures designed to avoid contamination from foreground stars and faint galaxies. These measurements were background subtracted using the mode of a rectangular image section southwest of clump 2. Because it was uncrowded, clump 2E was photometered with a 15-arcsec-radius circular aperture, implementing background subtraction in a surrounding annulus. On the basis of the surface density of sources comparable to clump 2E (in ultraviolet-optical colour and ultraviolet surface brightness) in the surrounding field, we estimate only a 4% chance of random association with the H I clump. For each region, we measured the flux (or determined an upper limit) in the FUV, NUV and r bands. Archival KPNO 4-m Mosaic data<sup>25</sup> was used for clumps 1 and 2, whereas Sloan Digital Sky Survey imaging<sup>26</sup> was employed for clump 2E. Table 1 lists selected properties of each analysed source. The areas over which we integrated contain multiple peaks (clumps 1 and 2) or extended structure (clump 2E) at GALEX resolution (200 pc), so we have not applied aperture corrections to the source magnitudes even for the clump-2E circular aperture. Values in Table 1 have been corrected for Galactic extinction. Dust in the Galactic foreground produces a minimal colour excess of  $E(B - V) = 0.024$  mag (ref. 27). In the GALEX bands, assuming a standard extinction curve<sup>28,29</sup>, this implies that  $A(\text{FUV}) = 0.19$  mag and  $A(\text{NUV}) = 0.21$  mag.

model/data offset was obtained for  $Z_\odot/50$  (red)– $Z_\odot/5$  (orange). At the assumed extinction, models for  $Z = 2.5Z_\odot$ ,  $Z_\odot$  (blue) and  $Z_\odot/2.5$  (green) are excluded by  $FUV - r$  colour. Metal-enriched models would also force inferred ages to be less than a crossing time.

In comparison with our models, the colours of clump 2 are anomalous. However, the distribution of ultraviolet colour within the clump-2 aperture is bimodal (Fig. 3). The source may suffer from significant NUV and minimal r-band contamination.

Thus, for simplicity, we attribute the ultraviolet emission in all clumps to continuous star formation at  $Z \approx Z_\odot/50$  over timescales of  $\sim 150$ – $550$  Myr for clumps 1 and 2E, and  $\sim 1$  Gyr for clump 2 (using only  $FUV - r$ ). If the stellar initial mass function is steeper or truncated at high mass in the ring clumps (as suspected in other low-surface-brightness environments<sup>16</sup>), these ages will be reduced because the models shift towards the red owing to a relative lack of FUV-producing stars. Additional observations are needed to address the issue of the initial mass function directly.

If star formation in the clumps was quasi-continuous on recent timescales, we can estimate the average integrated star formation rate

**Table 1 | Properties of selected Leo ring ultraviolet-bright complexes**

Property (unit)	Clump 1	Clump 2	Clump 2E
RA (J2000) (deg.)	161.9451	161.9789	162.0583
Dec. (J2000) (deg.)	12.1864	12.2369	12.0407
$ S_{\text{HI}} dv  (\text{Jy km s}^{-1})$	1.030	0.990	—
$M_{\text{HI}} (10^6 M_{\odot})$	35.2	33.9	—
$M_{\text{vir}} (10^6 M_{\odot})$	49.8	43.7	—
$v_{\text{hel}} (\text{km s}^{-1})$	987	956	—
Ultraviolet size ( $\text{kpc}^2$ )	4.1	2.4	1.7
FUV (ABmag)*	$20.37 \pm 0.08$	$20.62 \pm 0.08$	$20.72 \pm 0.07$
NUV (ABmag)	$20.57 \pm 0.09$	$20.58 \pm 0.07$	$21.03 \pm 0.12$
FUV – NUV (ABmag)	$-0.20 \pm 0.17$	$0.05 \pm 0.15$	$-0.31 \pm 0.19$
$r$ (ABmag)†	$>20.21 \pm 0.22$	$>20.10 \pm 0.15$	$>21.03 \pm 0.64$
$FUV - r$ (ABmag)†	$<0.16 \pm 0.30$	$<0.53 \pm 0.22$	$<-0.31 \pm 0.71$
$L_{\text{FUV},0} (10^{39} \text{ erg s}^{-1})$ ‡	9.0	7.1	6.5
Approximate age (Myr)§	550	1,000	150
SFR ( $M_{\odot} \text{ yr}^{-1}$ )	0.00077	0.00059	0.00063
$\Sigma_{\text{SFR}} (M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2})$ ¶	0.00019	0.00025	0.00037

The first six properties are taken from ref. 5, whereas the rest are derived from our GALEX/visible data set.  $|S_{\text{HI}} dv|$  is the total HI line flux after integrating over frequency ( $v$ ), and  $M_{\text{vir}}$  denotes the estimated virial mass,  $v_{\text{hel}}$  the measured heliocentric velocity and  $L_{\text{FUV},0}$  the extinction-corrected FUV luminosity.

\* All magnitudes are corrected only for Galactic extinction and quoted with  $1\sigma$  errors.

† The integrated  $r$ -band magnitudes should be viewed as upper limits on the flux until deep, sub-arcsecond imaging becomes available to rule out possible contamination from distant galaxies in our apertures. Consequently, the  $FUV - r$  values listed could also be marginally too red. Our  $r$ -band magnitudes are consistent with previous surface brightness limits (see ref. 2).

‡ Intrinsic FUV luminosity corrected for assumed internal extinction, with  $E(B - V) = 0.05$  mag.

§ Weakly constrained on the basis of point position in the colour–colour diagram of Fig. 4. In the case of clump 2, we used only  $FUV - r$  to estimate the age.

|| Computed assuming a continuous star formation epoch of the age specified in this table, and determined using population synthesis models.

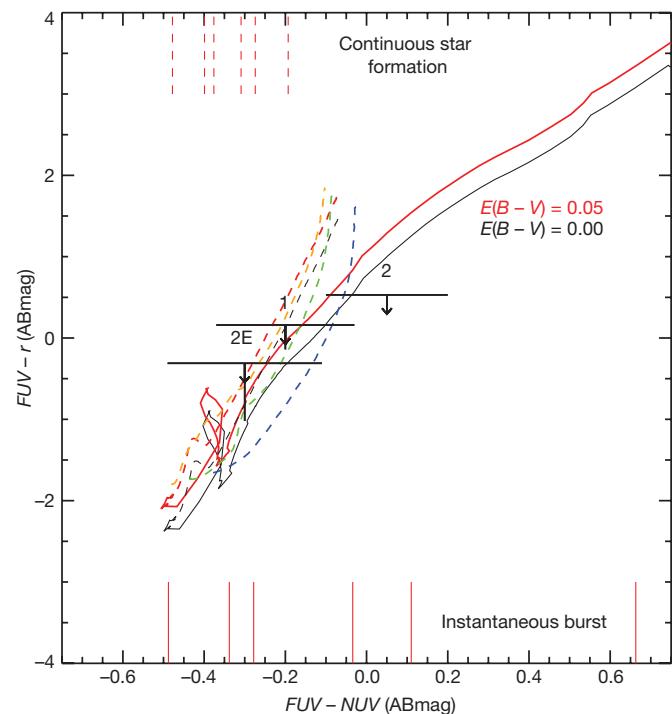
¶ Computed as SFR divided by ultraviolet size.

(SFR) for each region on the basis of the ultraviolet continuum luminosity (Table 1). Comparing observed and synthetic magnitudes for the FUV band, and assuming that  $E(B - V) = 0.05$  mag, we infer that  $\text{SFR} \approx (6-8) \times 10^{-3} M_{\odot} \text{ yr}^{-1}$  for the regions. We estimate SFR surface density ( $\Sigma_{\text{SFR}}$ ) values of  $(\sim 2-4) \times 10^{-4} M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ , comparable to azimuthally averaged  $\Sigma_{\text{SFR}}$  values measured in the outer disks of spiral galaxies<sup>17,18</sup>.

The southern Leo ring is forming stars. Perhaps this has also happened previously (or in other clumps) and escaped detection. Given the large ring age ( $>4 \times 10^9$  yr for at least one complete orbit), the probability of observing it at the epoch it began star formation is low, unless the star formation was prompted by interaction with M96. We are conducting a deep GALEX survey of the northern ring to look for recent star formation in clumps causally independent of M96. Ancient star formation episodes also remain a possibility. The visible sources associated with clumps 1 and 2 are inconspicuous and were not noted until our detection of ultraviolet light correlated with the H I concentrations. Other ring clouds could host intermittent star formation which remains unnoticed<sup>3,4</sup> because the most recent episode occurred over several hundred megayears ago, leaving no population of ultraviolet-producing stars. A  $10^6 M_{\odot}$  star formation event distributed inside a representative clump size of  $5 \text{ kpc}^2$  occurring 5 Gyr ago would have  $\mu_V \approx 28.8$  Vegamag arcsec $^{-2}$ , challenging the limit of modern surface photometry<sup>19</sup>. However, a deep, sub-arcsecond resolution imaging survey of the Leo ring to look for associated red-giant-branch stars could probe the primordial structure's long-term star formation history (and metallicity).

It is of interest to consider the eventual fate of the current star formation complexes. The characteristics (for example size and mass) of the Leo ring clumps are similar to those of dwarf galaxies<sup>5</sup>, and the objects appear to be self-gravitating<sup>5</sup>. Considering the detected star formation, we may reasonably label the clumps as dwarf galaxies in formation. It is instructive to compare the star formation clumps with blue compact dwarf prototype I Zwicky 18 (I Zw 18; ref. 20) and the transition dwarf Leo T (a ‘hobbit’ galaxy; ref. 21). The  $M_{\text{HI}}$  of clumps 1 and 2 is similar to that of I Zw 18, but an order of magnitude more than that of Leo T. However, the dynamical mass of I Zw 18 is substantially higher than those of the Ring substructures ( $2.6 \times 10^8 M_{\odot}$  versus  $4 \times 10^7 M_{\odot}$ ), owing to there being dark matter in I Zw 18 but not in the clumps<sup>5</sup>. The SFR of I Zw 18 is 50–100 times greater than in the clumps, perhaps owing to dark matter. Nevertheless, the recent SFR in Leo ring complexes is  $\sim 70$  times the average SFR in Leo T (refs 21, 22). Young stars in blue compact dwarfs and Leo T are complemented by an old stellar population<sup>20,22</sup>, but such evolved stars may be missing in the Leo clumps (unless they have evaded detection as we described). We suggest that the Leo ring clumps represent a ‘condensing’ stage in dwarf formation; however, these particular structures differ from ordinary halo-dominated dwarfs in that they lack a substantial dark matter component. In that sense, and considering their youth, the Leo ring complexes are similar to tidal dwarf galaxies<sup>23</sup>, although they form in a more pristine medium ( $Z_{\odot}/50$  versus  $Z_{\odot}/3$ ) not pre-enriched by a star-forming galaxy.

If gaseous structures such as the Leo ring were more common among galaxy groups in the early universe (before their expected ionization), and long-lived, dark-matter-free dwarfs can form within them, then it is possible that there exists an entire population of very faint, non-tidal, halo-lacking galaxies yet to be discovered. Within tidal tails, incipient substructures similar to tidal dwarf galaxies infrequently ( $\sim 25\%$ ) survive longer than 1 Gyr (ref. 23), as they can fall back onto the parent galaxy or experience tidal disruption from the parent. However, the ring environment is likely to be more amenable to survival of bound dwarfs, given that the giant clouds comprising it are comparatively isolated from massive galaxies near the group barycentre. The mode of formation for dwarf galaxies spawned in the primordial debris of group assembly, as witnessed in the Leo ring, makes them a ‘non-recycled’ product of cosmological structure



**Figure 4 | Colour–colour diagram comparing the Leo ring regions with population synthesis models<sup>14</sup>.** We show ( $FUV - NUV$ ,  $FUV - r$ ) predictions for instantaneous burst and continuous star formation scenarios. The black (thin) lines show colours for  $E(B - V) = 0$  mag whereas the red (thick) lines are for  $E(B - V) = 0.05$  mag, assuming  $Z_{\odot}/50$  in each case. The dashed and solid curves are continuous and, respectively, burst models. For the continuous star formation scenario, reddened case, we also plot predictions for  $Z_{\odot}/5$ ,  $Z_{\odot}/2.5$  and  $Z_{\odot}$  with thick, dashed orange, green and blue lines, respectively. Only the blueward  $1\sigma$  error bar is drawn for each  $FUV - r$  measurement, emphasizing that our values are upper limits. The ticks plotted on the top and bottom axes show ages of 2, 10, 20, 100, 200, and 1,000 Myr for each star formation scenario with  $E(B - V) = 0.05$  mag.

formation fundamentally different from ordinary, halo-dominant galaxies.

Received 3 November 2008; accepted 13 January 2009.

1. Schneider, S. E., Helou, G., Salpeter, E. E. & Terzian, Y. Discovery of a large intergalactic HI cloud in the M96 group. *Astrophys. J.* **273**, L1–L15 (1983).
2. Schneider, S. E. *et al.* Multifrequency survey of the intergalactic cloud in the M96 group. *Astron. J.* **97**, 666–673 (1989).
3. Pierce, M. J. & Tully, R. B. The extragalactic HI cloud in Leo. *Astron. J.* **90**, 450–453 (1985).
4. Kibblewhite, E. J., Dawson, M. G. M., Disney, M. J. & Phillips, S. An optical search for the intergalactic HI cloud in Leo. *Mon. Not. R. Astron. Soc.* **213**, 111–115 (1985).
5. Schneider, S. E., Salpeter, E. E. & Terzian, Y. High-resolution observations of the intergalactic HI cloud in the M96 group. *Astron. J.* **91**, 13–22 (1986).
6. Sil'chenko, O. K., Moiseev, A. V., Afanasiev, V. L., Chavushyan, V. H. & Valdes, J. R. The Leo I cloud: Secular nuclear evolution of NGC 3379, NGC 3384, and NGC3368? *Astrophys. J.* **591**, 185–203 (2003).
7. Rood, H. J. & Williams, B. A. The intergalactic HI cloud in Leo – a simple modelling of the Spitzer-Baade collision event. *Astrophys. J.* **288**, 535–550 (1985).
8. Spitzer, L. Jr & Baade, W. Stellar populations and collisions of galaxies. *Astrophys. J.* **113**, 413–418 (1951).
9. Bekki, K., Koribalski, B. S., Ryder, S. D. & Couch, W. J. Massive HI clouds with no optical counterparts as high-density regions of intragroup HI rings and arcs. *Mon. Not. R. Astron. Soc.* **357**, L21–L25 (2005).
10. Schneider, S. E. Neutral hydrogen in the M96 group – the galaxies and the intergalactic ring. *Astrophys. J.* **343**, 94–106 (1989).
11. Reynolds, R. J., Magee, K., Roesler, F. L., Scherb, F. & Harlander, J. H-alpha scans of the intergalactic HI cloud in Leo. *Astrophys. J.* **309**, L9–L12 (1986).
12. Donahue, M., Aldering, G. & Stocke, J. T. Low surface brightness H-alpha observations of local intergalactic hydrogen clouds. *Astrophys. J.* **450**, L45–L49 (1995).
13. Martin, D. C. *et al.* The Galaxy Evolution Explorer: a space ultraviolet survey mission. *Astrophys. J.* **619**, L1–L6 (2005).
14. Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).
15. Gil de Paz, A. *et al.* Chemical and photometric evolution of extended ultraviolet disks: optical spectroscopy of M83 (NGC 5236) and NGC 4625. *Astrophys. J.* **661**, 115–134 (2007).
16. Meurer, G. *et al.* Evidence for a non-uniform initial mass function in the local Universe. *Astrophys. J.* (in the press).
17. Thilker, D. A. *et al.* A search for extended ultraviolet disk (XUV-disk) galaxies in the local universe. *Astrophys. J. Suppl. Ser.* **173**, 538–571 (2007).
18. Boissier, S. *et al.* Radial variation of attenuation and star formation in the largest late-type disks observed with GALEX. *Astrophys. J. Suppl. Ser.* **173**, 524–537 (2007).
19. Mihos, J. C., Harding, P., Feldmeier, J. & Morrison, H. Diffuse light in the Virgo Cluster. *Astrophys. J.* **631**, L41–L44 (2005).
20. Aloisi, A. *et al.* I Zw 18 revisited with HST ACS and Cepheids: new distance and age. *Astrophys. J.* **667**, L151–L154 (2007).
21. Ryan-Weber, E. V. *et al.* The local group dwarf Leo T: HI on the brink of star formation. *Mon. Not. R. Astron. Soc.* **384**, 535–540 (2008).
22. de Jong, J. T. A. *et al.* The structural properties and star formation history of Leo T from deep LBT photometry. *Astrophys. J.* **680**, 1112–1119 (2008).
23. Bournaud, F. & Duc, P.-A. From tidal dwarf galaxies to satellite galaxies. *Astron. Astrophys.* **456**, 481–492 (2006).
24. Schneider, S. Neutral hydrogen in the M96 group – evidence for a giant intergalactic ring. *Astrophys. J.* **288**, L33–L35 (1985).
25. Millis, R. L. *et al.* The Deep Ecliptic Survey: a search for Kuiper Belt Objects and Centaurs. I. Description of methods and initial results. *Astron. J.* **123**, 2083–2109 (2002).
26. Adelman-McCarthy, J. K. *et al.* The sixth data release of the Sloan Digital Sky Survey. *Astrophys. J. Suppl. Ser.* **175**, 297–313 (2008).
27. Schlegel, D. J., Finkbeiner, D. P. & Davis, M. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *Astrophys. J.* **500**, 525–553 (1998).
28. Cardelli, J. A., Clayton, G. C. & Mathis, J. S. The relationship between infrared, optical, and ultraviolet extinction. *Astrophys. J.* **345**, 245–256 (1989).
29. O'Donnell, J. E.  $R_V$ -dependent optical and near-ultraviolet extinction. *Astrophys. J.* **422**, 158–163 (1994).

**Acknowledgements** GALEX is a NASA Small Explorer, launched in April 2003. We gratefully acknowledge NASA's support of the construction, operation, and science analysis for the GALEX mission, developed in cooperation with the Centre National d'Etudes Spatiales, France, and the Korean Ministry of Science and Technology. This research draws upon data provided by B. Millis as distributed by the National Optical Astronomy Observatory (NOAO) Science Archive. NOAO is operated by the Association of Universities for Research in Astronomy under a cooperative agreement with the National Science Foundation. This research has made use of the NASA/IPAC Extragalactic Database, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. This research made use of NASA's Astrophysics Data System.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to D.A.T. (dthilker@pha.jhu.edu).

## LETTERS

## Travelling-wave nuclear magnetic resonance

David O. Brunner<sup>1</sup>, Nicola De Zanche<sup>1</sup>, Jürg Fröhlich<sup>2</sup>, Jan Paska<sup>2</sup> & Klaas P. Pruessmann<sup>1</sup>

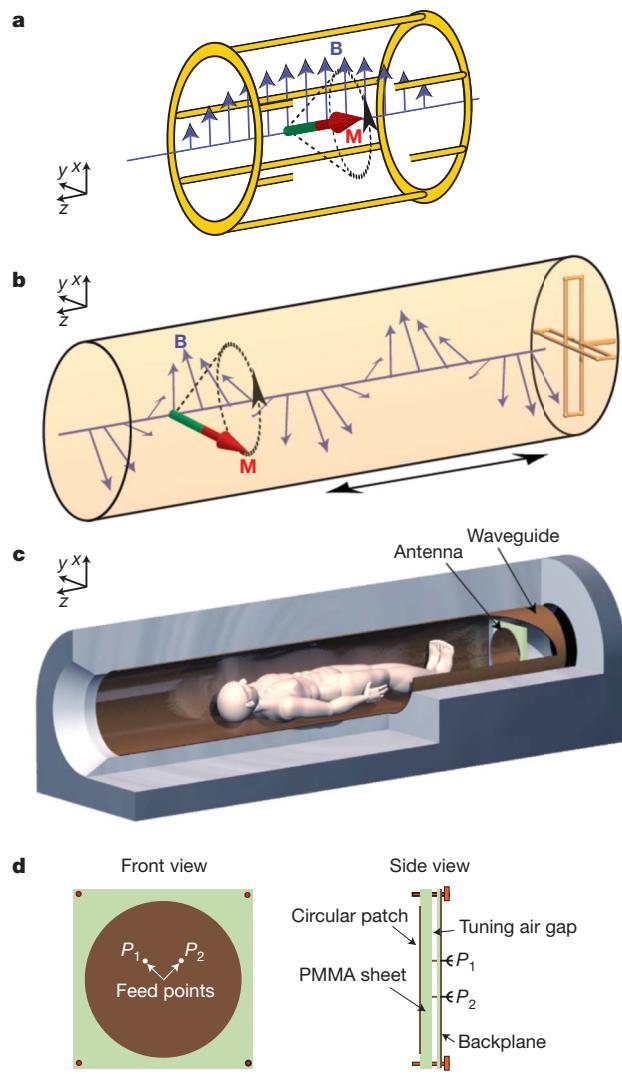
Nuclear magnetic resonance<sup>1,2</sup> (NMR) is one of the most versatile experimental methods in chemistry, physics and biology<sup>3</sup>, providing insight into the structure and dynamics of matter at the molecular scale. Its imaging variant—magnetic resonance imaging<sup>4,5</sup> (MRI)—is widely used to examine the anatomy, physiology and metabolism of the human body. NMR signal detection is traditionally based on Faraday induction<sup>6</sup> in one or multiple radio-frequency resonators<sup>7–10</sup> that are brought into close proximity with the sample. Alternative principles involving structured-material flux guides<sup>11</sup>, superconducting quantum interference devices<sup>12</sup>, atomic magnetometers<sup>13</sup>, Hall probes<sup>14</sup> or magnetoresistive elements<sup>15</sup> have been explored. However, a common feature of all NMR implementations until now is that they rely on close coupling between the detector and the object under investigation. Here we show that NMR can also be excited and detected by long-range interaction, relying on travelling radio-frequency waves sent and received by an antenna. One benefit of this approach is more uniform coverage of samples that are larger than the wavelength of the NMR signal—an important current issue in MRI of humans at very high magnetic fields. By allowing a significant distance between the probe and the sample, travelling-wave interaction also introduces new possibilities in the design of NMR experiments and systems.

Uniform spatial coverage in NMR and MRI is traditionally achieved by tailoring the reactive near field of resonant Faraday probes<sup>7–10</sup>. This approach is valid when the radio-frequency wavelength at the Larmor frequency is substantially larger than the target volume, which does not hold for modern, wide-bore, high-field systems. At the highest field strength currently used for human studies, 9.4 T (refs 16, 17), the resonance frequency of hydrogen nuclei reaches 400 MHz, corresponding to a wavelength in tissue on the order of 10 cm. At such short wavelengths, head or body resonators form standing-wave field patterns, which degrade MRI results by causing regional signal losses and perturbing the contrast between different types of tissue.

The non-uniformity of standing waves is due to the underlying electrodynamics, which require that the magnetic field exhibit curvature according to its frequency and the ambient material. Standing waves fulfil this condition by having spatial variation in the field magnitude (Fig. 1a). However, the required field curvature can also be translated, partly or wholly, into phase variation. By causing the underlying field pattern to propagate through space, such phase variation reduces the variation of the field magnitude. Notably, the limiting case of a plane wave has a perfectly uniform magnitude at any wavelength. In addition, travelling radio-frequency waves offer a natural means of exciting and detecting NMR across large distances (Fig. 1b).

Despite these attractive features, travelling-wave NMR has not been explored so far. In traditional cylindrical set-ups, the formation of travelling waves at the NMR frequency is suppressed by structures surrounding the sample, such as gradient coils, cryostats, and radio-frequency screens. Their conductive surfaces admit axially travelling waves only beyond some cut-off frequency that is roughly inversely

proportional to the bore width. Therefore, travelling-wave NMR requires a high-field magnet that also has a wide bore to bring the cut-off frequency below the NMR frequency.



**Figure 1 | Working principles of traditional and travelling-wave NMR.** **a**, Traditional resonant probes form a standing radio-frequency wave within the sample. Its magnetic component, **B**, causes nutation of the nuclear magnetization, **M**, and governs the probe's receive sensitivity. **b**, In our approach, an antenna probe interacts with the sample through a travelling wave. **c**, In a wide-bore, high-field magnet, such waves can be guided by a conductive lining, permitting remote NMR excitation and detection with an antenna at the end of the magnet. **d**, Sketch of the circularly polarized patch antenna used for the initial implementation of this idea. PMMA, poly(methyl methacrylate).

<sup>1</sup>Institute for Biomedical Engineering, University of Zürich and ETH Zürich, Gloriastrasse 35, 8092 Zürich, Switzerland. <sup>2</sup>Laboratory for Electromagnetic Fields and Microwave Electronics, ETH Zürich, Gloriastrasse 35, 8092 Zürich, Switzerland.

To fulfil this requirement, we used a cylindrical, superconducting 7.0-T magnet with a 58-cm-diameter bore lined with a radio-frequency screen (Fig. 1c). The screen was made from a stainless-steel mesh designed to provide high conductance at radio frequencies while blocking audio-frequency eddy currents induced by the surrounding gradient coils. When enclosing only air, the bore has a cut-off frequency of 303 MHz, which is still just above the proton Larmor frequency of 298 MHz. However, the frequency limit is reduced when dielectric material is brought into the bore. Even small amounts of dielectric loading enable the formation of axially travelling waves at the NMR frequency, effectively using the radio-frequency screen as a waveguide. A human body in particular—containing large amounts of water, which is a strong dielectric—reduces the cut-off frequency sufficiently to clearly enter the travelling-wave regime.

NMR using such travelling waves requires a new type of probe. Instead of coupling to the reactive near field of the sample, a travelling-wave probe must couple to the propagating modes of the waveguide. To do so, it no longer needs to be close to the sample but can be placed anywhere along the bore. Requirements of this sort are well known in microwave engineering and can be addressed by a range of technical solutions. For the present work, the NMR probe was implemented in the form of a circularly polarized patch antenna<sup>18</sup> (Fig. 1d).

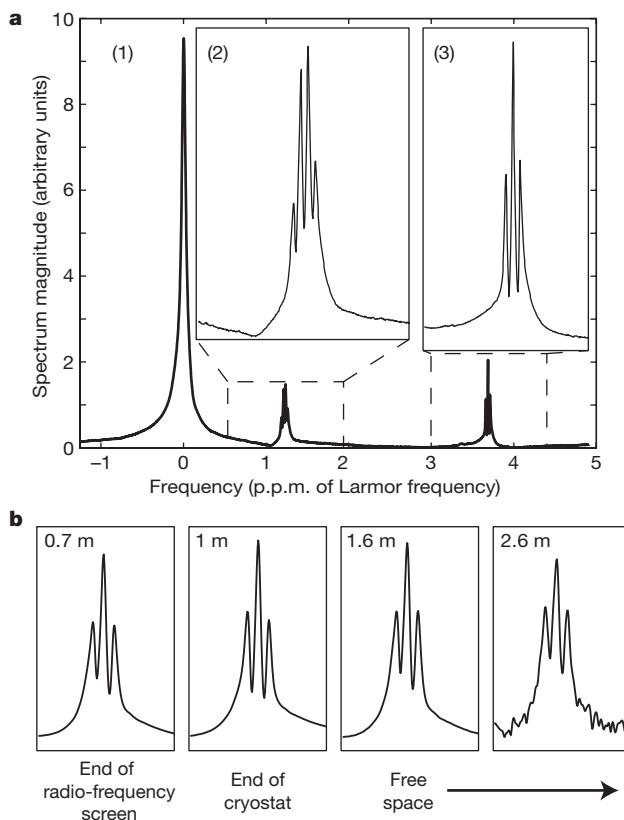
Using this set-up, the principle of travelling-wave NMR was first demonstrated by spectroscopy of an aqueous 10% ethanol solution. Proton NMR was excited and detected by the patch antenna, which was initially mounted at the end of the radio-frequency screen, 70 cm from the sample. The resulting spectrum (Fig. 2a) shows the expected dominant water peak as well as the methyl and methylene resonances of the ethanol molecule. The experiment was then repeated with gradually increasing antenna distances. As the magnified methyl triplets in Fig. 2b show, consistent spectrum quality was obtained with the probe placed

well outside the magnet and a well-resolved spectrum was still detected at a distance of 2.6 m from the sample. The evident loss of sensitivity at large distances reflects the expected decrease in coupling between the antenna and the modes of the bore. Higher sensitivity at large distances would be achieved using an antenna of greater directivity or a longer waveguide.

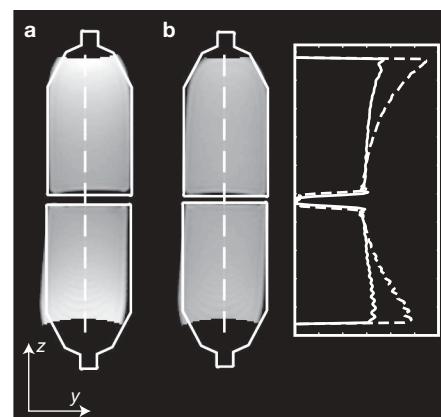
The spatial uniformity of the travelling waves was studied in an extended sample 50 cm in length, formed by two adjacent bottles filled with mineral oil. On its own, this arrangement did not yield completely uniform radio-frequency coverage, as illustrated by the MRI results shown in Fig. 3a. The residual non-uniformity indicates the presence of a standing radio-frequency wave superimposed on the intended travelling component. It is caused mainly by slight reflections at the transitions between the bottles and the empty sections of the bore, which entail changes in wave impedance. The reflections can be mitigated by wave impedance matching and additional loading. To demonstrate this, two further bottles were added towards the open end of the bore, the second one containing a conductive water solution to act as a termination dissipating incident wave energy. This modification rendered the MRI results substantially more uniform, indicating the presence of almost purely travelling radio-frequency waves (Fig. 3b).

For reasons of safety, initial *in vivo* experiments targeted only a volunteer's lower extremities, ensuring that the chest and head remained outside the waveguide. The antenna was placed at the opposite end of the bore, 70 cm from the ankle. The resulting magnetic resonance image (Fig. 4a) shows the right lower leg with good uniformity over a large volume. The field of view of 50 cm is the maximum possible with the gradient system used and is not limited by the travelling-wave concept. Wave impedance matching was not necessary in this case, because the leg per se forms a sufficiently smooth, tapered impedance transition. For comparison, the same imaging procedure was repeated with a commercially available birdcage resonator optimized for head MRI at 7 T. The result thus obtained (Fig. 4b) shows smaller coverage, reflecting inherent limitations of resonant probes. The standing-wave nature of its rung currents limits the feasible length of the birdcage probe, which is 17 cm for this model. On longer rungs, the radio-frequency current would become even more non-uniform<sup>19</sup>, causing a similar axial sensitivity drop-off and rendering the resonator unstable<sup>20</sup>.

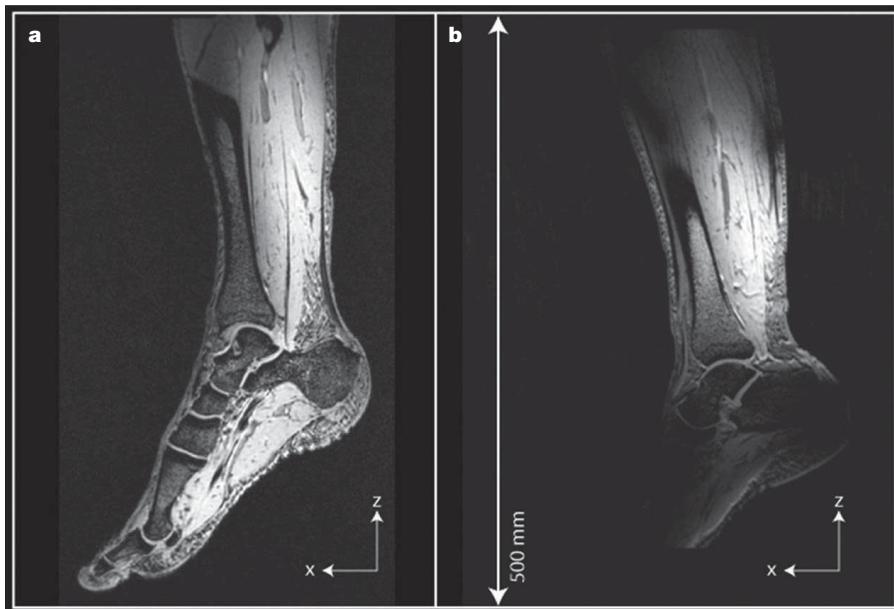
Although it was possible to cover the lower legs uniformly, significant radio-frequency attenuation is expected to occur along the full length of a human body. Simulations of the set-up in Fig. 1c, assuming an adult male subject, indicate that in total the body absorbs approximately 90% of the radio-frequency power coupled into the waveguide and only the remaining 10% are radiated from its far end. Less attenuation is expected for shorter or slimmer subjects or when



**Figure 2 | Demonstration of travelling-wave NMR in an aqueous 10% ethanol solution.** a, NMR spectrum obtained at an antenna distance of 70 cm, showing water (1), methylene (2), and methyl (3) resonances. Centre frequency of experiment, 298.04 MHz. b, Details of the methyl triplet as observed with increasing distance, as indicated.



**Figure 3 | Example of wave impedance matching in travelling-wave MRI.** a, Non-uniform coverage of two phantom bottles is caused by residual standing radio-frequency waves, as also shown by the dashed image intensity profile on the right. b, The standing waves can be suppressed using wave impedance matching and dissipation in a termination load (solid profile).



**Figure 4 | In vivo results.** **a**, Travelling-wave MRI of a human lower leg *in vivo*. **b**, Identical scan performed with a traditional resonant probe.

using an even wider bore. On the basis of the same simulations, the coupling efficiency of the patch antenna was estimated at 80%, a value that can be improved by optimizing the strategy of driving the waveguide modes.

Apart from coverage and uniformity, the travelling-wave concept will also affect the sensitivity and radio-frequency power efficiency of NMR experiments. According to the reciprocity of NMR signals<sup>21</sup>, the sensitivity of a radio-frequency probe is closely related to its efficiency, that is, its yield of circularly polarized radio-frequency magnetic field at reference input power. Resonant near-field probes achieve high efficiency by concentrating radio-frequency energy and dissipation in the sample and the probe itself. By contrast, the travelling-wave approach relies on radio-frequency energy flowing through the set-up, requiring that part of it be absorbed beyond the target volume. This can be done using a dedicated absorber device or diffusely outside the waveguide as was the case in our initial experiments. With either solution, the necessary absorber losses will take some toll in terms of efficiency and sensitivity, constituting a drawback of travelling-wave probes in comparison with resonators.

Reduced probe efficiency is the lesser concern, because it can be addressed by using higher driving power. *In vivo* MRI is usually not limited by technical radio-frequency power constraints but rather by sample heating, to which the absorber losses do not contribute. The corresponding sensitivity loss, caused by thermal noise originating from any material that absorbs radio-frequency power during transmission, is more limiting. One potential way of avoiding this loss is to use a cryogenically cooled absorber structure. For MRI applications, it is also conceivable to combine travelling-wave excitation with local detection by an array of detunable surface resonators<sup>9</sup>. Such a hybrid approach will reconcile the improved coverage and safety advantage of travelling-wave excitation with the sensitivity benefit of close-range array detection.

With respect to net sensitivity, another potential concern is the phase delay that results from signal propagation to and from the resonant nuclei. For large samples, it will give rise to significant phase differences between signals travelling different distances. This effect is illustrated in Fig. 5a, which shows travelling-wave MRI of a long, water-filled cylinder with antennas placed at both ends of the waveguide. Using the same antenna for transmission and reception results in a linear distribution of the image phase. MRI will not be hampered by such phase variation as long as it is resolved by the imaging process. However, spectroscopic experiments could suffer from delay-related

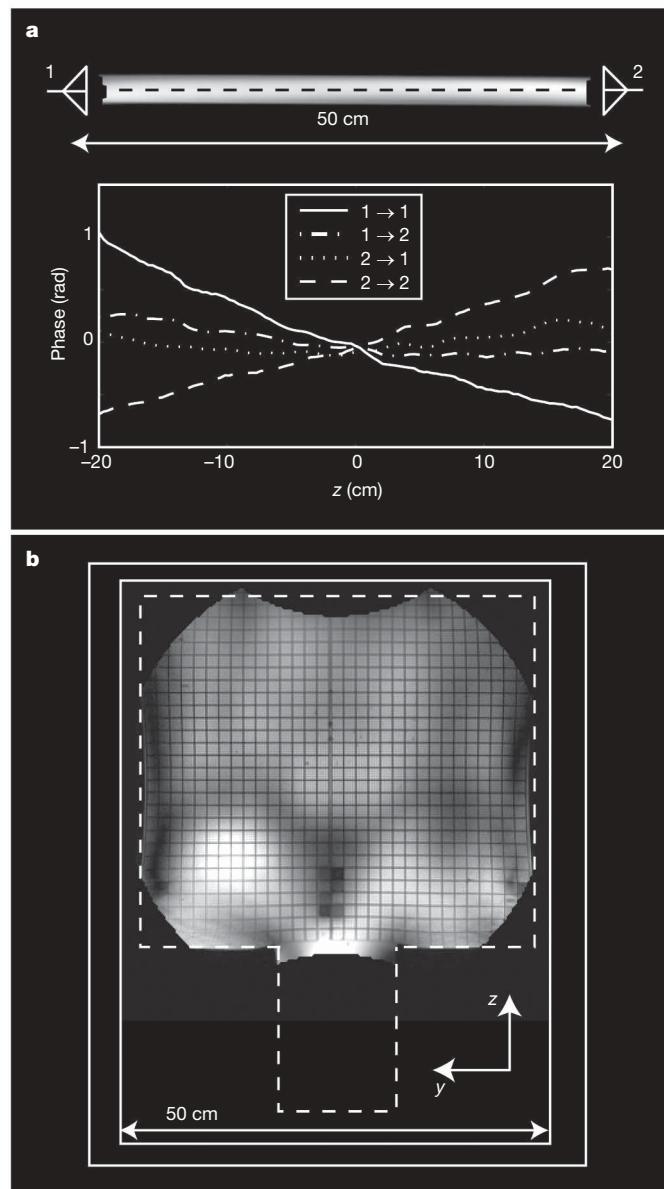
dephasing. To address this problem, the transmit and receive antennas should in general be designed and positioned such that the total phase delay is the same across the volume of interest. In the abovementioned experiment, this was accomplished by using either of the two antennas for transmission and the other for reception. As shown in Fig. 5a, the linear phase patterns nearly vanished in these configurations. Alternatively, after travelling-wave excitation from either side, pulsed gradient encoding could be applied to compensate for variable phase delay, effectively refocusing the spin radiation for reception by the same antenna as used for radio-frequency transmission.

As it relies on far-field electrodynamics, travelling-wave NMR prompts analogies with electron spin resonance<sup>22</sup> and nonlinear optics<sup>23</sup>. Adopting principles and devices from these fields may enable the study of phenomena analogous to, for example, photon echoes<sup>24</sup>, four-wave mixing<sup>25</sup> and self-induced transparency<sup>26</sup>. Potentially useful analogies can also be drawn with the large variety of more widespread technologies that rely on travelling-wave phenomena, including manifold imaging modalities, radar and telecommunication. In the present work, the established concepts of waveguides and antennas have already proven useful.

By adopting propagating waves, NMR also incurs known complications of this phenomenon, such as material-dependent diffraction and attenuation. These effects will be encountered, for example, in MRI of the human torso, which has pronounced dielectric interfaces such as the shoulders. Figure 5b shows an initial experimental illustration of dielectric perturbations in a phantom of similar size and shape as a human torso. One way of countering these effects is wave impedance matching, which has been demonstrated here in a basic form. Analogously to index matching in microwave and optical technology, wave impedance matching can be achieved using a wide range of distributions of dielectric, magnetically permeable or conductive material and thus offers great inherent freedom for tailoring resulting radio-frequency field distributions.

By improving the extent and uniformity of spatial coverage, travelling-wave MRI has the potential to facilitate the exploration of the highest field strengths currently available for human studies. A further promising area of application is high-field screening. The ability to perform spatially resolved NMR in larger volumes may simplify studies of large numbers of small animals<sup>27</sup> or inanimate samples in parallel.

Finally, introducing a significant distance between the sample and the NMR probe has multiple beneficial side effects. In such a configuration, the probe is not loaded by losses in the sample, which simplifies



**Figure 5 | Travelling-wave MRI of very large samples.** **a**, Two antennas, one at each end of the waveguide, were used to image a long water cylinder. Transmission and reception with the same antenna ( $1 \rightarrow 1$ ,  $2 \rightarrow 2$ ) give rise to linear phase delays, as shown in graphs of the image phase along the phantom. Constant net phase delay is achieved by transmitting with either of the antennas and receiving with the opposite one ( $1 \rightarrow 2$ ,  $2 \rightarrow 1$ ). **b**, Image of a human torso phantom. Dielectric boundaries introduce discontinuities of the wave impedance, causing field perturbations that diminish towards the lower regions of the phantom.

impedance matching and renders the probe performance substantially more robust than that of near-field high-frequency probes. This situation also simplifies safety considerations in human studies by avoiding exposure to strong, short-range electric fields emanating from the probe. Furthermore, placing the probe far away frees space in the centre of costly high-field magnets, may improve the comfort of human subjects and facilitates the positioning of alternative equipment, such as stimulation devices for studies of brain function.

## METHODS SUMMARY

All experiments were performed using a 7-T whole-human-body MRI system (Philips Healthcare). The cylindrical radio-frequency screen (diameter, 58 cm) was made from a stainless-steel mesh. The patch antenna<sup>18</sup> consisted of a thin copper disk and a backplane, each on PMMA formers and was driven in quadrature.

The spectra shown in Fig. 2 were obtained using a stimulated-echo acquisition mode (STEAM) sequence for volume selection. The sample was an aqueous 10% ethanol solution in a 1-l glass beaker and was axially tapered by bottles filled with distilled water. The data shown in Fig. 3 were acquired with a two-dimensional gradient-echo sequence, using low flip angles to represent the radio-frequency field distribution faithfully. Each bottle was filled with 3 l of Marcol 82 mineral oil. The *in vivo* data of Fig. 4 were acquired using a low-flip-angle, three-dimensional gradient-echo sequence with identical sequence parameters for both scans. Only the radio-frequency power was reduced to 25% in the resonator scan to implement similar flip angles in the centre. The transfer and attenuation of radio-frequency power in this set-up were estimated by finite-difference time-domain (FDTD) simulation incorporating the antenna, the waveguide and a detailed anatomical model of an adult male human (IT'IS Foundation)<sup>28</sup>. For the study reported in Fig. 5a, two-dimensional gradient-echo images were acquired from a 2-m-long, water-filled cylinder 7 cm in diameter. Two antennas were used simultaneously, one at each end of the waveguide. In two successive sets of experiments, either of the antennas was used for transmission and image signals were received with both. Performing all scans with two different echo times permitted subtraction of the phase effect of local resonance offsets. The torso phantom in Fig. 5b was filled with 40 l of gelled water with conductivity similar to human tissue.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 25 September; accepted 24 December 2008.

- Bloch, F., Hansen, W. W. & Packard, M. The nuclear induction experiment. *Phys. Rev.* **70**, 474–485 (1946).
- Purcell, E. M., Torrey, H. C. & Pound, R. V. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.* **69**, 37–38 (1946).
- de Graaf, R. A. *NMR Spectroscopy* (Wiley, 2007).
- Lauterbur, P. C. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature* **242**, 190–197 (1973).
- Kumar, A., Welti, D. & Ernst, R. R. NMR Fourier zeugmatography. *J. Magn. Reson.* **18**, 69–83 (1975).
- Hahn, E. L. Nuclear induction due to free Larmor precession. *Phys. Rev.* **77**, 297–298 (1950).
- Hayes, C. E., Edelstein, W. A., Schenck, J. F., Mueller, O. M. & Eash, M. An efficient, highly homogeneous radiofrequency coil for whole-body NMR imaging at 1.5 T. *J. Magn. Reson.* **63**, 622–628 (1985).
- Tropp, J. Theory of the birdcage resonator. *J. Magn. Reson.* **82**, 51–62 (1989).
- Roemer, P. B., Edelstein, W. A., Hayes, C. E., Souza, S. P. & Mueller, O. M. The NMR phased array. *Magn. Reson. Med.* **16**, 192–225 (1990).
- Vaughan, J. T., Hetherington, H. P., Otu, J. O., Pan, J. W. & Pohost, G. M. High frequency volume coils for clinical NMR imaging and spectroscopy. *Magn. Reson. Med.* **32**, 206–218 (1994).
- Wiltshire, M. C. K. *et al.* Microstructured magnetic materials for RF flux guides in magnetic resonance imaging. *Science* **291**, 849–851 (2001).
- Day, E. P. Detection of NMR using a Josephson-junction magnetometer. *Phys. Rev. Lett.* **29**, 540–542 (1972).
- Savukov, I. M. & Romalis, M. V. NMR detection with an atomic magnetometer. *Phys. Rev. Lett.* **94**, 123001 (2005).
- Boero, G., Besse, P.-A. & Popovic, R. Hall detection of magnetic resonance. *Appl. Phys. Lett.* **79**, 14948–1501 (2001).
- Verpillat, F. *et al.* Remote detection of nuclear magnetic resonance with an anisotropic magnetoresistive sensor. *Proc. Natl. Acad. Sci. USA* **105**, 2271–2273 (2008).
- Vaughan, T. *et al.* 9.4T human MRI: Preliminary results. *Magn. Reson. Med.* **56**, 1274–1282 (2006).
- Atkinson, I. C., Renteria, L., Burd, H., Pliskin, N. H. & Thulborn, K. R. Safety of human MRI at static fields above the FDA 8T guideline: Sodium imaging at 9.4T does not affect vital signs or cognitive ability. *J. Magn. Reson. Imaging* **26**, 1222–1227 (2007).
- Balanis, C. A. *Antenna Theory: Analysis and Design* 2nd edn, Ch. 14 (Wiley, 2005).
- Bogdanov, G. & Ludwig, R. Coupled microstrip line transverse electromagnetic resonator model for high-field magnetic resonance imaging. *Magn. Reson. Med.* **47**, 579–593 (2002).
- Harpen, M. D. Cylindrical coils near self-resonance. *Magn. Reson. Med.* **30**, 489–493 (1993).
- Hoult, D. I. The principle of reciprocity in signal strength calculations: a mathematical guide. *Concepts Magn. Reson.* **12**, 173–187 (2000).
- Gulla, A. F. & Budil, D. E. Engineering and design concepts for quasi-optical high-field electron paramagnetic resonance. *Concepts Magn. Reson. B* **22B**, 15–36 (2004).
- Boyd, R. W. *Nonlinear Optics* (Elsevier Science, 2003).
- Kurnit, N. A., Abella, I. D. & Hartmann, S. R. Observation of a photon echo. *Phys. Rev. Lett.* **13**, 567–568 (1964).
- Armstrong, J. A., Bloembergen, N., Ducuing, J. & Pershan, P. S. Light waves at the boundary of nonlinear media. *Phys. Rev.* **128**, 606–622 (1962).

26. McCall, S. L. & Hahn, E. L. Self-induced transparency. *Phys. Rev.* **183**, 457–490 (1969).
27. Bock, N. A., Konyer, N. B. & Henkelman, R. M. Multiple-mouse MRI. *Magn. Reson. Med.* **49**, 158–167 (2003).
28. Christ, A. et al. Development of CAD based anatomical human body models of two adults and two children. *EBAE 2007*, abstr. S-4-2 (8th Internat. Congr. Eur. BioElectromag. Assoc., 2007).

**Acknowledgements** We thank N. van den Berg and A. Trabesinger for discussions. We are also grateful to P. Boesiger for his leading role in creating the 7T facility. This work was funded in part by the Swiss National Science Foundation (Project 116400) and by the Velux Foundation. Technical support from Philips Healthcare is also gratefully acknowledged.

**Author Contributions** D.O.B.: basic concept, antenna design and construction, bench experiments, magnetic resonance experiments, manuscript. N.D.Z.: conceptual considerations, assistance with antenna design, assistance with bench and magnetic resonance experiments, editing. J.F.: conceptual considerations, FDTD models, radio-frequency safety validation, editing. J.P.: FDTD models. K.P.P.: conceptual considerations, assistance with magnetic resonance experiments, manuscript, supervision.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to K.P.P. ([pruessmann@biomed.ee.ethz.ch](mailto:pruessmann@biomed.ee.ethz.ch)).

## METHODS

All experiments were performed in a cylindrical superconducting 7.0-T magnet (Philips Healthcare) equipped with a three-axis gradient system and lined with a radio-frequency screen made from a stainless-steel mesh (diameter, 58 cm; length, 135 cm). NMR and MRI data acquisition was performed with an integrated console and spectrometer (Achieva, Philips Healthcare). Throughout, the radio-frequency transmitter power was limited to the same values as for human head exposure with a local resonator, that is, 10 W average forward power and 1.4 kW peak power.

The custom-designed patch antenna is shown in Fig. 1d. Its front plane is a 350-mm-diameter disk made of copper sheet glued onto a 25-mm-thick PMMA former. The backplane is formed by a square copper sheet mounted on 5-mm-thick PMMA. PMMA yields negligible proton NMR signal at the echo times used in this work. A variable gap between the two planes was used to tune the antenna to the proton NMR frequency of 298 MHz. The antenna was driven at two points in the front plane forming a right angle with the centre of the disk. The radial positions of the feed points were chosen such as to match the impedance of the 50- $\Omega$  feed lines. To produce circular polarization in the transmission mode, the two ports were fed through a 90° hybrid splitter. In receive operation, the two channels were connected to independent receive lines of the spectrometer and combined digitally. Switching between transmit and receive operation was performed by diode switches.

The spectra shown in Fig. 2 were obtained from a 10% solution of ethanol in water, contained in a 1-l glass beaker. To support the formation of axially travelling waves, bottles of distilled water were added on either side of the beaker as dielectric loads. NMR spectra were acquired with a localized STEAM sequence with an echo time of 11 ms and a repetition time of 3 s. Low-flip-angle radio-frequency pulses were used such that the selected volume of 9 cm<sup>3</sup> changed negligibly with the radio-frequency amplitude. The acquisition bandwidth was 2 kHz and 16 phase cycles were performed to cancel spurious echoes. When the position of the radio-frequency probe was changed, all console and spectrometer settings were kept constant.

The data shown in Fig. 3 were obtained from bottles filled with Marcel 82 mineral oil, using a gradient-echo sequence with an echo time of 2.9 ms, a repetition time of 42 ms and a bandwidth per pixel of 338 Hz. A very small flip angle was chosen to ensure that the image intensity depended linearly on the transmit and receive sensitivities of the probe.

The *in vivo* images of a volunteer's right lower leg (Fig. 4) were acquired with a three-dimensional gradient-echo sequence, yielding an isotropic resolution of 1 mm in 6 min. The echo time was 3.1 ms, the repetition time was 12 ms and the bandwidth per pixel was 376 Hz. All sequence parameters were kept the same for the two probes, except for the transmit power, which was reduced to 25% for the resonator to obtain similar flip angles in the centre. The axial length of the field of

view of 500 mm was again limited by the gradient system. The resonant probe was a commercially available birdcage resonator with a diameter of 30 cm and a length of 17 cm (Nova Medical). Both legs were enclosed within the resonator to keep the volunteer in the same position.

Electromagnetic simulations of the *in vivo* setup were performed using the FDTD technique (SEMCAD, Schmid & Partner Engineering AG). The cylindrical radio-frequency screen was modelled as a perfect electrical conductor and an anatomical model of an adult male<sup>28</sup>, featuring *in vivo* dielectric tissue properties, was positioned at its centre. The patch antenna was placed as in the *in vivo* experiment, that is, at a distance of 70 cm from the ankle. The energy flow of radio-frequency waves through the waveguide was assessed by integrating the Poynting vector over transverse control planes. A first control plane, between the antenna and the model human, determined the power coupled into the waveguide. A second control plane, in the loss-free region close to the end of the waveguide, was used to calculate the power radiated from the far end of the waveguide. The power difference was verified by volume integration of ohmic losses over the entire body.

In the experiments reported in Fig. 5a, a radio-frequency-spoiled, two-dimensional gradient-echo sequence was used to image a 2-m-long, water-filled cylinder 7 cm in diameter. Two antennas were used simultaneously, one at each end of the bore, and were placed at distances of 10 cm from the ends of the waveguide to minimize reflections. In these experiments, only a single port of each antenna was used for radio-frequency transmission and signal reception, to avoid any phase perturbation by imperfect port combination. In two successive sets of experiments, either of the antennas was used for transmission and image signals were received with both. All scans were performed twice with an echo time difference of 0.5 ms, permitting the estimation and removal of any phase contributions from local frequency offsets. The remaining image phase results exclusively from radio-frequency propagation, including unknown constant offsets of various sources, such as cabling and spectrometer calibrations. To eliminate these arbitrary offsets, each phase plot was offset corrected to null the phase in the centre ( $z = 0$ ). The nominal field of view was 50 cm. To avoid phase distortions due to potential gradient delays, the read-out direction was chosen to be orthogonal to the cylinder axis and no correction for gradient nonlinearity was applied. The range of meaningful phase values was thereby reduced to approximately 35 cm. The diameter of the cylinder appears reduced because the imaged slice was slightly off-centre.

Figure 5b shows a torso phantom made of PMMA as proposed by the American Society for Testing and Materials, filled with gelled water (relative permittivity,  $\epsilon_r \approx 80$ ; conductivity,  $\sigma \approx 0.47 \text{ S m}^{-1}$ ). Imaging of this phantom was performed by a low-flip-angle, multiple-slice, radio-frequency-spoiled, two-dimensional gradient-echo sequence with a repetition time of 50 ms per slice, yielding an effective in-plane resolution of 1 mm and a slice thickness of 4 mm.

## LETTERS

# Magnetic assembly of colloidal superstructures with multipole symmetry

Randall M. Erb<sup>1\*</sup>, Hui S. Son<sup>1\*</sup>, Bappaditya Samanta<sup>2</sup>, Vincent M. Rotello<sup>2</sup> & Benjamin B. Yellen<sup>1</sup>

The assembly of complex structures out of simple colloidal building blocks is of practical interest for building materials with unique optical properties (for example photonic crystals<sup>1</sup> and DNA biosensors<sup>2</sup>) and is of fundamental importance in improving our understanding of self-assembly processes occurring on molecular to macroscopic length scales<sup>3–5</sup>. Here we demonstrate a self-assembly principle that is capable of organizing a diverse set of colloidal particles into highly reproducible, rotationally symmetric arrangements. The structures are assembled using the magnetostatic interaction between effectively diamagnetic and paramagnetic particles within a magnetized ferrofluid. The resulting multipolar geometries resemble electrostatic charge configurations such as axial quadrupoles ('Saturn rings'), axial octupoles ('flowers'), linear quadrupoles (poles) and mixed multipole arrangements ('two tone'), which represent just a few examples of the type of structure that can be built using this technique.

Colloidal particle dispersions feature a wide variety of attractive and repulsive interactions and mechanical packing constraints that have been exploited to assemble a diverse set of structures, including Pickering emulsions and particle clusters<sup>4,6–8</sup>, linear and kinked chains<sup>9–11</sup>, and highly ordered arrays<sup>12,13</sup> and matrices<sup>14,15</sup>. Self-assembly mediated by Van der Waals forces and hydrophobic interactions leads primarily to close-packed structures such as colloidal crystals when the assembly is performed in the bulk fluid<sup>1,16,17</sup>, and to particle clusters when the assembly process is confined to the interior of an emulsion droplet or to a liquid–liquid interface<sup>4,6,8</sup>. Ionic interactions between oppositely charged colloidal particles have been used to assemble various ionic crystalline lattices<sup>14,15</sup>. Also, both electric and magnetic dipole–dipole interactions have been used to create either one-dimensional chain structures or two-dimensional arrays when attractive or repulsive interactions, respectively, are induced in the particles by an external field<sup>9,11,18,19</sup>. Despite recent progress in colloidal self-assembly, there have been relatively few demonstrations of ordered structures in suspensions containing multiple different particle types. At present, the demonstrations of controlled assembly in binary suspensions include ionic crystalline lattices<sup>14,15</sup> and discrete particle clusters assembled in oil-in-water emulsions<sup>3,20</sup>; however, there have been no demonstrations of controlled assembly in suspensions containing three or more colloidal components.

Here we demonstrate the use of dipole–dipole interactions to fashion multi-component, rotationally symmetric colloidal superstructures from isotropic spherical particles. We produce these structures by applying static, uniform magnetic fields to aqueous suspensions of paramagnetic and non-magnetic particles suspended in a ferrofluid (that is, Fe<sub>3</sub>O<sub>4</sub> nanoparticles suspended in water). The ferrofluid concentration is an essential control parameter for this assembly technique. Particles with higher magnetization than their carrier fluid (for example paramagnetic particles) will exhibit a classical paramagnetic response with

respect to the surrounding fluid, that is, with the particle moments aligned parallel to the external field direction. Conversely, particles less magnetizable than the carrier fluid (for example non-magnetic particles) will behave as non-magnetic cavities inside a magnetizable medium<sup>21</sup> and will exhibit an effectively diamagnetic response with respect to the surrounding fluid, that is, with the particle moments aligned antiparallel to the external field direction. In multi-component suspensions, the paramagnetic and non-magnetic particles interact antiferromagnetically, resulting in the formation of ring structures (Fig. 1) that differ from the chain and sheet assemblies commonly observed when electric or magnetic fields are applied to single-component colloidal suspensions<sup>9,11,18,19,22</sup>.

A physical model can be used to predict the equilibrium structure produced within multi-component systems. This model employs a continuum approach to approximate the local fluid magnetization, which is justified when the ferrofluid particles are much smaller than other colloidal particles in the fluid. The effective dipole moment of linearly magnetizable spherical particles immersed in a magnetizable carrier fluid is given by

$$\mathbf{m}_i = 4\pi a^3 \frac{\chi_i - \chi_f}{\chi_i + 2\chi_f + 3} \mathbf{H}_{\text{ext}} \quad (1)$$

where  $a$  is the radius of the particle,  $\mathbf{H}_{\text{ext}}$  is the magnetic field at the location of the particle centre and  $\chi_f$  and  $\chi_i$  are the magnetic susceptibilities of the ferrofluid and the particle, respectively<sup>21</sup>. If an appropriate ferrofluid is chosen such that the susceptibility of the fluid is between that of the non-magnetic and paramagnetic particles (that is,  $\chi_{\text{NM}} < \chi_f < \chi_{\text{PM}}$ ), then equation (1) predicts that the paramagnetic particles, with  $\chi_i = \chi_{\text{PM}} > \chi_f$  will behave as point dipoles aligned parallel to the applied field, whereas the non-magnetic particles, with  $\chi_i = \chi_{\text{NM}} \approx 0$ , will behave as point dipoles aligned antiparallel to the field. Although 'non-magnetic' materials typically have slight intrinsic paramagnetic or diamagnetic properties with susceptibility values on the order of  $10^{-6}$ , all of these materials will have effectively negligible susceptibilities in comparison with a ferrofluid (typical susceptibilities of ferrofluids are on the order of 0.1–1.0), and their effective moments will therefore reflect the model proposed here by assuming that  $\chi_{\text{NM}} \approx 0$ .

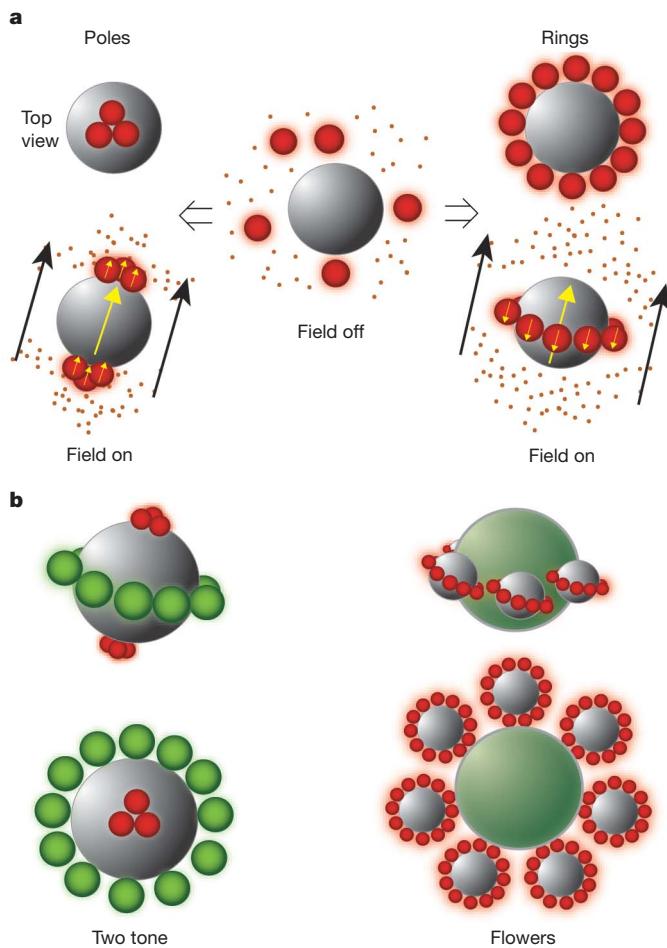
Stability of particle configurations can be understood from the expression for magnetostatic potential energy of a dipole  $\mathbf{m}$  in a locally applied field  $\mathbf{H}$ , given by

$$U_i(\mathbf{r}) = -2\pi a^3 \mu_0 \frac{\chi_i - \chi_f}{\chi_i + 2\chi_f + 3} |\mathbf{H}(\mathbf{r})|^2 \quad (2)$$

where  $\mu_0$  is the magnetic permeability of free space. For the non-magnetic particle ( $\chi_{\text{NM}} < \chi_f$ ), equation (2) implies that its potential energy is strictly positive, which causes the particle to move towards regions of minimum magnetic field,  $|\mathbf{H}(\mathbf{r})| = |\mathbf{H}_{\text{min}}|$ . Equation (2)

<sup>1</sup>Duke University, Department of Mechanical Engineering and Materials Science, Center for Biologically Inspired Materials and Material Systems, Box 90300, Hudson Hall, Durham, North Carolina 27708, USA. <sup>2</sup>Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003, USA.

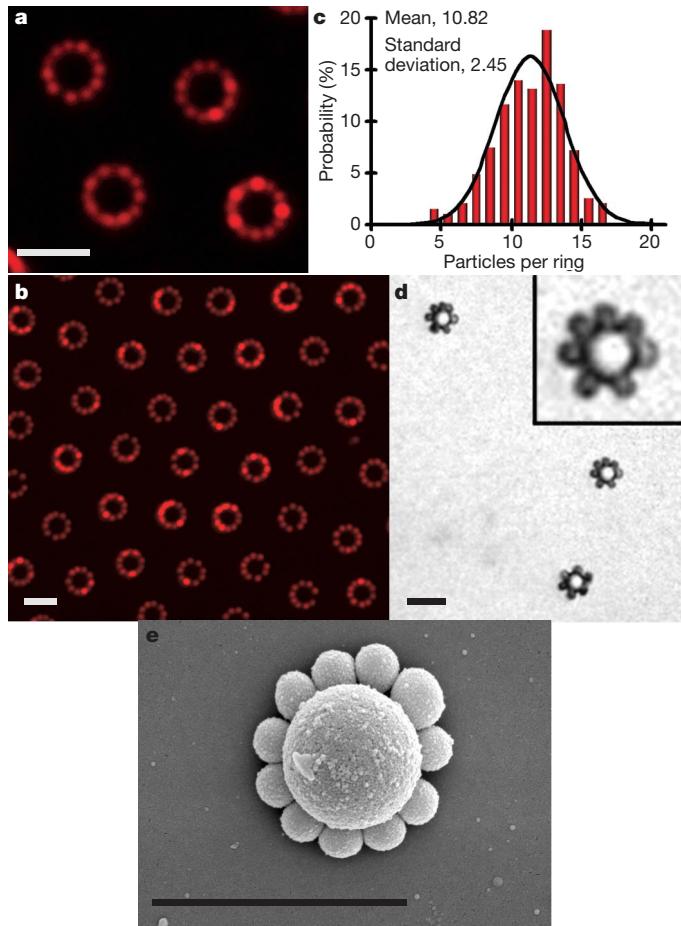
\*These authors contributed equally to this work.



**Figure 1 | Illustration of magnetic assembly in colloidal particle mixtures.** **a**, Rotationally symmetric structures form in mixtures of ferrofluid (brown dots) and small/large colloidal particles (red/grey spheres, respectively). No structures form in the absence of a magnetic field. When a magnetic field is applied (black arrows denote the external field direction), equatorial and polar arrangements form when the particle magnetizations respectively have the opposite and the same sense. Particle dipole moments are indicated by the yellow arrows. **b**, Hierarchical assembly in three-component suspensions is illustrated. ‘Two-tone’ structures consist of simultaneous polar and equatorial arrangements assembling on a single core particle, whereas ‘flower’ structures are formed by the fractal assembly of multiple equatorial arrangements.

also implies that the paramagnetic particles ( $\chi_{PM} > \chi_f$ ) have strictly negative potential energy, causing them to move towards regions of maximum field,  $|\mathbf{H}(\mathbf{r})| = |\mathbf{H}_{max}|$ . Hence, in mixed suspensions of non-magnetic and paramagnetic particles, magnetic force drives the non-magnetic particle towards the ‘equator’ of the paramagnetic particle, where the dipolar field of the paramagnetic particle subtracts from the external field. Likewise, the paramagnetic particles move towards the equator of the non-magnetic particles, where the dipolar field of the non-magnetic particle adds to the external field. Under the assumption that stable structures can be formed only when the magnetostatic potential energy greatly exceeds thermal fluctuations, we calculate that the smallest particle that can be manipulated using this technique is on the order of 60 nm in diameter. (See calculations in Supplementary Information and ref. 23).

Experimental demonstration of this assembly technique is presented in Fig. 2. The size ratio between the colloidal particles is an important control parameter used to adjust the resulting structures. Large paramagnetic particles mixed with smaller non-magnetic particles will provide ring configurations (Fig. 2a, b), where the paramagnetic particle acts as the magnetic core around which the satellite non-magnetic



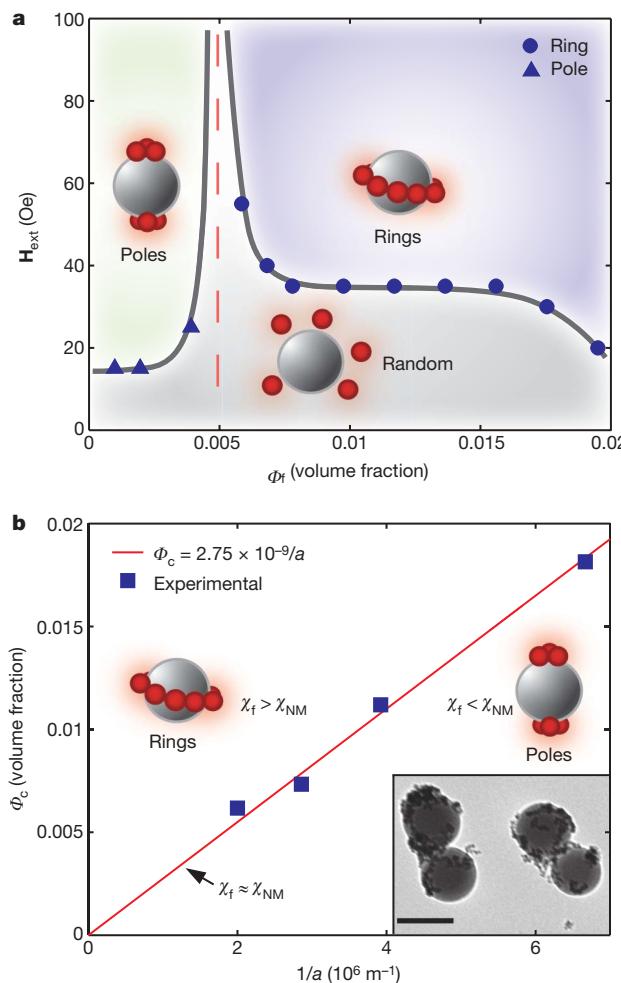
**Figure 2 | Assembly of ‘Saturn-rings’ particles and their statistical distribution.** **a, b**, Rings form in a three-component aqueous suspension including ferrofluid (1.2% volume fraction,  $\chi_f = 0.25$ ), 2.7-μm paramagnetic particles (0.33% volume fraction) and 1.0-μm (red) non-magnetic particles (0.2% volume fraction), corresponding to 11–12 non-magnetic particles per paramagnetic particle. **c**, An average of 10.82 non-magnetic particles per ring confirms this dilution ratio. **d**, Inverse rings form inside a suspension of ferrofluid (2.2% volume fraction,  $\chi_f = 0.46$ ), 3.0-μm non-magnetic particles (0.32% volume fraction) and 1.0-μm paramagnetic particles (0.057% volume fraction), corresponding to 5 paramagnetic particles per non-magnetic particle. Inset, magnified view of one of the inverse rings. **e**, SEM image showing ring structure in dry state. Scale bars, 5 μm.

particles assemble in a Saturn-rings configuration. The inverse structure is formed when large non-magnetic particles are mixed with smaller paramagnetic particles (Fig. 2d). The formation of ring shapes was verified both by confocal microscopy and by observing the synchronous rotation of the colloidal assemblies in a rotating magnetic field (Supplementary Video 1). The assemblies form within seconds of the application of an external field, and dissolve when the external field is removed, indicating the rapid and switchable nature of the assembly technique (Supplementary Video 2).

A notable aspect of this magnetic assembly technique is its self-limiting nature. The magnetostatic interaction between two nearby colloidal assemblies is dominated by the repulsive force between the core particles within each assembly, resulting in large separations between individual structures (Fig. 2b). This self-limiting process can assist in the formation of highly monodisperse colloidal structures of controlled shape by permanently linking the structures using an appropriate cross-linking strategy. We have demonstrated that the ring structures can be stabilized using either streptavidin–biotin recognition (Supplementary Video 3) or by covalent linking of amine-functionalized particles with glutaraldehyde as the cross-linking agent by previously reported methods<sup>11</sup>. The resulting structures have been purified using a combination of centrifugation and magnetic

separation steps; a scanning electron microscope (SEM) image of a colloidal ring in the dry state is provided as an example (Fig. 2e). These ring structures are highly stable when suspended in fluid even after multiple separation and rinsing steps; however, we have observed that the strong evaporative force causes the satellite non-magnetic beads to reorganize around the central magnetic bead during the drying step (Supplementary Fig. 3).

Additional colloidal structures can be engineered by adjusting the degree of the non-magnetic particle's magnetization. Our ferrofluid was designed to controllably adsorb on the surfaces of non-magnetic particles, as confirmed in transmission electron microscopy (TEM) measurements (Fig. 3). In very dilute ferrofluid suspensions, the non-magnetic particles with an adsorbed layer of  $\text{Fe}_3\text{O}_4$  nanoparticles have higher magnetization than the surrounding medium and are thus attracted to the poles of the paramagnetic particles instead of to the equator. The critical concentration for the pole–ring transition can be determined from simple arguments based on the magnetization difference between the non-magnetic particle (including its adsorbed  $\text{Fe}_3\text{O}_4$  nanoparticle layer) and the surrounding fluid, approximated as



**Figure 3 | Experimental phase diagrams and critical behaviour.**

**a**, Experimental phase diagram depicting the dependence of ferrofluid concentration,  $\Phi_f$ , on the evolution of structural phases in suspensions of 1.0- $\mu\text{m}$  non-magnetic particles (0.25% volume fraction) and 2.7- $\mu\text{m}$  paramagnetic particles (0.25% volume fraction).  $H_{\text{ext}}$  is the critical field strength of the phase transition. **b**, The critical ferrofluid concentration,  $\Phi_c$ , implying susceptibility matching between the non-magnetic particle and the surrounding ferrofluid, is plotted as a function of the inverse particle diameter. Inset, TEM image of  $\text{Fe}_3\text{O}_4$  nanoparticles adsorbed onto polystyrene beads. Scale bar, 200 nm.

$$\mathbf{M} = (4\pi a^2 \tau \theta \chi_B / V - \chi_f) \mathbf{H}_{\text{ext}} \quad (3)$$

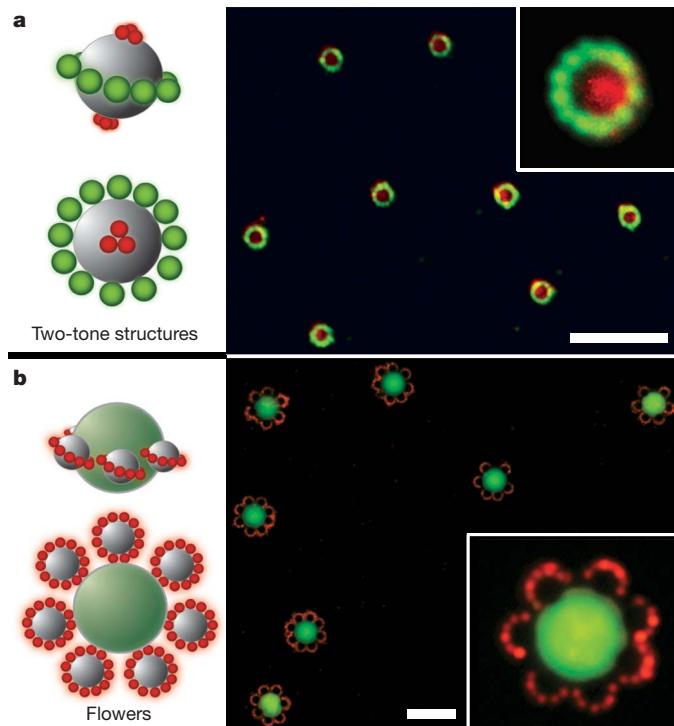
where  $a$ ,  $\tau$ ,  $\theta$ ,  $\chi_B$  and  $V$  are the radius of the non-magnetic particle, the diameter of the  $\text{Fe}_3\text{O}_4$  nanoparticle, the fractional area coverage of  $\text{Fe}_3\text{O}_4$  nanoparticles adsorbed on the non-magnetic particle, the bulk susceptibility of undiluted ferrofluid and the volume of the non-magnetic particle, respectively. Magnetic susceptibility measurements indicate that within the tested concentration range the susceptibility of the ferrofluid follows a linear relationship with the  $\text{Fe}_3\text{O}_4$  nanoparticle concentration, namely  $\chi_f = \chi_B \Phi_f$ , where  $\Phi_f$  is the volume fraction of  $\text{Fe}_3\text{O}_4$  nanoparticles and  $\chi_B$  is the extrapolated value for the susceptibility of bulk  $\text{Fe}_3\text{O}_4$  ( $\chi_B \approx 21$ ; Supplementary Fig. 1). The critical concentration,  $\Phi_c$ , for the pole–ring transition is determined by setting equation (3) to zero, leading to the following expression:

$$\Phi_c = 3\tau\theta/a \quad (4)$$

The critical concentration for each non-magnetic particle size was determined experimentally from the singularity in the phase diagram (Fig. 3a). Near the singularity, the field strength required to form colloidal structures increases dramatically as a result of approaching the index-matching condition between the non-magnetic particle and the surrounding fluid. A discussion of the analytic technique used to identify the phase transition is provided in the Supplementary Information (Supplementary Fig. 2 and related discussion). The inverse relationship between particle radius and critical ferrofluid concentration predicted by equation (4) is supported by experiments conducted with different-sized non-magnetic particles ranging from 300 nm to 1000 nm in diameter (Fig. 3b). Ring structures form above the best-fit line, where the non-magnetic bead's magnetic susceptibility,  $\chi_{NM} = \chi_B \Phi_c$ , is less than the ferrofluid susceptibility,  $\chi_f$ . Pole structures form below the best-fit line, where the converse is true. The slope suggests that the area fraction of  $\text{Fe}_3\text{O}_4$  nanoparticle coverage is  $\sim 9.2\%$ , assuming a uniformly distributed monolayer of 10-nm  $\text{Fe}_3\text{O}_4$  nanoparticles adsorbed to the non-magnetic particle's surface. TEM images (Fig. 3b, inset) reveal the presence of particles in the predicted locations, with aggregation that may be either inherent in the system or a result of drying. Visual inspection of the number of adsorbed nanoparticles on the non-magnetic particle's surface suggests a reasonable agreement with our calculated value of fractional area coverage.

More complex structures can be formed using this magnetic assembly technique. We present two initial possibilities here. By assembling both pole and ring configurations simultaneously, two-tone structures can be formed. This is accomplished by exploiting the sensitivity of the particle's size to the pole–ring transition. The two-tone structures (Fig. 4a) are assembled in a mixture of 2.7- $\mu\text{m}$ -diameter paramagnetic particles, 1.0- $\mu\text{m}$ -diameter (green) non-magnetic particles, 0.21- $\mu\text{m}$ -diameter (red) non-magnetic particles and ferrofluid at an appropriate dilution (in the range of  $\sim 1\%$  volume fraction, corresponding to magnetic susceptibility of  $\chi_f \approx 0.2$ ). In these conditions, the smaller non-magnetic particles acquire positive magnetization with respect to the ferrofluid, whereas the larger non-magnetic particles acquire negative magnetization with respect to the ferrofluid. Hence, the smaller particles are attracted to the poles of the paramagnetic particle, whereas the larger particles are attracted to the equator of the same paramagnetic particle (Supplementary Video 4).

Other types of complex hierarchical structure can be formed using the ring structures as building blocks for assembling larger superstructures. The flower-shaped colloidal structures (Fig. 4b) are formed when 9.9- $\mu\text{m}$  (fluorescent green) non-magnetic particles are mixed with 2.7- $\mu\text{m}$  paramagnetic particles and 1.0- $\mu\text{m}$  (red) non-magnetic particles (Supplementary Video 5). The flowers in Fig. 4b are notable for the high uniformity in the number of 'petals' in each, which shows that the assembled structures can be obtained in high yield, typically with 6–8 petals per flower at these relative concentrations. Qualitatively, the number of petals was observed to



**Figure 4 | Demonstration of multi-component particle assembly.**

**a, b**, Fluorescent images of structure formation in four-component colloidal-particle aqueous suspensions consisting of ferrofluid (1.1% volume fraction,  $\chi_f = 0.24$ ), 0.21- $\mu\text{m}$  (red) non-magnetic particles (0.02% volume fraction), 1.0- $\mu\text{m}$  (green) non-magnetic particles (0.2% volume fraction) and 2.7- $\mu\text{m}$  paramagnetic core particles (0.33% volume fraction) (**a**) and ferrofluid (1.2% volume fraction,  $\chi_f = 0.25$ ), 1.0- $\mu\text{m}$  (red) non-magnetic particles (0.12% volume fraction), 2.7- $\mu\text{m}$  paramagnetic core particles (0.21% volume fraction) and 9.9- $\mu\text{m}$  (green) non-magnetic particles (1.5% volume fraction) (**b**). Each panel includes a sketch of the assembled structure and a magnified view of one of the assembled structures (inset). Scale bars, 20  $\mu\text{m}$ .

increase at higher ferrofluid concentrations and decrease at lower ferrofluid concentrations, as a result of the diminishing or, respectively, increasing influence of dipole–dipole repulsion between the magnetic beads of each pair of petals.

In conclusion, application of magnetically actuated self-assembly techniques to multi-particle colloidal mixtures provides a direct and general approach to the creation of complex colloidal superstructures. A rich variety of different particle configurations is possible depending on the size, type and degree of magnetization of the different particle components. In this work, we have demonstrated that nonmagnetic particles as small as 200 nm in diameter can be assembled into complex arrangements; however, theoretical predictions and preliminary experimental tests have shown that particles smaller than this can be manipulated using other magnetic particles and ferrofluids<sup>23</sup>.

## METHODS SUMMARY

**Colloidal assembly protocols.** Colloidal assembly experiments were observed in a 5- $\mu\text{l}$  bead mixture placed between a coverslip and a glass slide and then sealed with microscope oil to reduce fluid motion. A uniform magnetic field was applied to the bead solution by passing current through air-core solenoids (Fisher Scientific) fitted with iron cores. Microscopy was performed with a DM LM fluorescent microscope (Leica) using  $\times 40$  air-immersion and  $\times 100$  oil-immersion (numerical aperture, 1.3) objectives and a dual red-green filter cube (Chroma Technology).

**Measurements and analysis.** The magnetic susceptibility of the ferrofluid was measured with a MPMS-7 superconducting quantum interference device magnetometer (Quantum Design). Critical thresholds for ring and pole formation

were defined by fluorescence intensity measurements as a function of the external field strength. For each field strength, the average intensity in the immediate vicinity of 25–30 individual colloidal assemblies was measured. The phase transition was defined as the maximum in the slope of the fluorescence intensity measured as a function of magnetic field.

**Crosslinking approach.** Streptavidin–biotin crosslinking was performed using 0.86- $\mu\text{m}$  biotin-coated non-magnetic beads (Spherotech) and 2.7- $\mu\text{m}$  streptavidin-coated paramagnetic beads (Dynal Biotech) rinsed in  $\times 0.1$  PBS. A solution was prepared with 1.9% ferrofluid, 0.32% paramagnetic beads and 0.2% non-magnetic beads by volume fraction. After 5–10 min exposure in a 100-Oe magnetic field, the non-magnetic beads were permanently crosslinked to the paramagnetic beads in ring structures. Glutaraldehyde crosslinking was performed using the above method except with 0.86- $\mu\text{m}$  amine-coated non-magnetic beads (Sigma-Aldrich) at 0.2% volume fraction in a 5% glutaraldehyde solution (MP Biomedicals) at pH 8–9. After field exposure, the solution was centrifuged to remove the iron oxide nanoparticles, then magnetically separated to remove spare non-magnetic beads, leaving purified rings that were dried and observed in the SEM.

Received 21 February 2008; accepted 7 January 2009.

- Xia, Y. N., Gates, B. & Li, Z. Y. Self-assembly approaches to three-dimensional photonic crystals. *Adv. Mater.* **13**, 409–413 (2001).
- Elghanian, R., Storhoff, J. J., Mucic, R. C., Letsinger, R. L. & Mirkin, C. A. Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. *Science* **277**, 1078–1081 (1997).
- Zerrouki, D., Baudry, J., Pine, D., Chaikin, P. & Bibette, J. Chiral colloidal clusters. *Nature* **455**, 380–382 (2008).
- Manoharan, V. N., Elsesser, M. T. & Pine, D. J. Dense packing and symmetry in small clusters of microspheres. *Science* **301**, 483–487 (2003).
- Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418–2421 (2002).
- Dinsmore, A. D. et al. Colloidosomes: selectively permeable capsules composed of colloidal particles. *Science* **298**, 1006–1009 (2002).
- Velev, O. D., Lenhoff, A. M. & Kaler, E. W. A class of microstructured particles through colloidal crystallization. *Science* **287**, 2240–2243 (2000).
- Kim, S. H., Yi, G. R., Kim, K. H. & Yang, S. M. Photocurable Pickering emulsion for colloidal particles with structural complexity. *Langmuir* **24**, 2365–2371 (2008).
- Dreyfus, R. et al. Microscopic artificial swimmers. *Nature* **437**, 862–865 (2005).
- Musevic, I., Skarabot, M., Tkalec, U., Ravnik, M. & Zumer, S. Two-dimensional nematic colloidal crystals self-assembled by topological defects. *Science* **313**, 954–958 (2006).
- Biswal, S. L. & Gast, A. P. Rotational dynamics of semiflexible paramagnetic particle chains. *Phys. Rev. E* **69**, 041406 (2004).
- Yellen, B. B., Hovorka, O. & Friedman, G. Arranging matter by magnetic nanoparticle assemblers. *Proc. Natl. Acad. Sci. USA* **102**, 8860–8864 (2005).
- Yin, Y., Lu, Y., Gates, B. & Xia, Y. Template-assisted self-assembly: A practical route to complex aggregates of monodispersed colloids with well-defined sizes, shapes, and structures. *J. Am. Chem. Soc.* **123**, 8718–8729 (2001).
- Shevchenko, E. V., Talapin, D. V., Kotov, N. A., O'Brien, S. & Murray, C. B. Structural diversity in binary nanoparticle superlattices. *Nature* **439**, 55–59 (2006).
- Leunissen, M. E. et al. Ionic colloidal crystals of oppositely charged particles. *Nature* **437**, 235–240 (2005).
- Ozin, G. A. & Yang, S. M. The race for the photonic chip: Colloidal crystal assembly in silicon wafers. *Adv. Funct. Mater.* **11**, 95–104 (2001).
- Zhuang, J., Wu, H., Yang, Y. & Cao, Y. C. Controlling colloidal superparticle growth through solvophobic interactions. *Angew. Chem. Int. Ed.* **47**, 2208–2212 (2008).
- Skjeltorp, A. T. One- and two-dimensional crystallization of magnetic holes. *Phys. Rev. Lett.* **51**, 2306–2309 (1983).
- Halsey, T. C. Electrorheological fluids. *Science* **258**, 761–766 (1992).
- Cho, Y. S. et al. Self-organization of bidisperse colloids in water droplets. *J. Am. Chem. Soc.* **127**, 15968–15975 (2005).
- Panofsky, W. K. H. & Phillips, M. *Classical Electricity and Magnetism* 84–86 (Addison Wesley, 1955).
- Helgesen, G., Pieranski, P. & Skjeltorp, A. T. Dynamic behavior of simple magnetic hole systems. *Phys. Rev. A* **42**, 7271–7280 (1990).
- Erb, R. M. & Yellen, B. B. Concentration gradients in mixed magnetic and nonmagnetic colloidal suspensions. *J. Appl. Phys.* **103**, 07A312 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors would like to thank the US National Science Foundation (grants NSEC DMR-0531171 and MRSEC DMR-0213695 to V.M.R. and grants CMMI-0608819 and CMMI-0625480 to B.B.Y.) for supporting this work.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to B.B.Y. (yellen@duke.edu).

## LETTERS

# Increasing carbon storage in intact African tropical forests

Simon L. Lewis<sup>1</sup>, Gabriela Lopez-Gonzalez<sup>1</sup>, Bonaventure Sonké<sup>2</sup>, Kofi Affum-Baffoe<sup>3</sup>, Timothy R. Baker<sup>1</sup>, Lucas O. Ojo<sup>4</sup>, Oliver L. Phillips<sup>1</sup>, Jan M. Reitsma<sup>5</sup>, Lee White<sup>6</sup>, James A. Comiskey<sup>7†</sup>, Marie-Noël Djukouo K.<sup>2</sup>, Corneille E. N. Ewango<sup>8</sup>, Ted R. Feldpausch<sup>1</sup>, Alan C. Hamilton<sup>9</sup>, Manuel Gloor<sup>1</sup>, Terese Hart<sup>10</sup>, Annette Hladik<sup>11</sup>, Jon Lloyd<sup>1</sup>, Jon C. Lovett<sup>12</sup>, Jean-Remy Makana<sup>10</sup>, Yadvinder Malhi<sup>13</sup>, Frank M. Mbago<sup>14</sup>, Henry J. Ndangalasi<sup>14</sup>, Julie Peacock<sup>1</sup>, Kelvin S.-H. Peh<sup>1</sup>, Douglas Sheil<sup>15†</sup>, Terry Sunderland<sup>7†</sup>, Michael D. Swaine<sup>16</sup>, James Taplin<sup>12</sup>, David Taylor<sup>17</sup>, Sean C. Thomas<sup>18</sup>, Raymond Votere<sup>3</sup> & Hannsjörg Wöll<sup>19</sup>

The response of terrestrial vegetation to a globally changing environment is central to predictions of future levels of atmospheric carbon dioxide<sup>1,2</sup>. The role of tropical forests is critical because they are carbon-dense and highly productive<sup>3,4</sup>. Inventory plots across Amazonia show that old-growth forests have increased in carbon storage over recent decades<sup>5–7</sup>, but the response of one-third of the world's tropical forests in Africa<sup>8</sup> is largely unknown owing to an absence of spatially extensive observation networks<sup>9,10</sup>. Here we report data from a ten-country network of long-term monitoring plots in African tropical forests. We find that across 79 plots (163 ha) above-ground carbon storage in live trees increased by  $0.63 \text{ Mg Ch}^{-1} \text{ yr}^{-1}$  between 1968 and 2007 (95% confidence interval (CI), 0.22–0.94; mean interval, 1987–96). Extrapolation to unmeasured forest components (live roots, small trees, necromass) and scaling to the continent implies a total increase in carbon storage in African tropical forest trees of  $0.34 \text{ Pg Cyr}^{-1}$  (CI, 0.15–0.43). These reported changes in carbon storage are similar to those reported for Amazonian forests per unit area<sup>6,7</sup>, providing evidence that increasing carbon storage in old-growth forests is a pan-tropical phenomenon. Indeed, combining all standardized inventory data from this study and from tropical America and Asia<sup>5,6,11</sup> together yields a comparable figure of  $0.49 \text{ Mg Ch}^{-1} \text{ yr}^{-1}$  ( $n = 156$ ; 562 ha; CI, 0.29–0.66; mean interval, 1987–97). This indicates a carbon sink of  $1.3 \text{ Pg Cyr}^{-1}$  (CI, 0.8–1.6) across all tropical forests during recent decades. Taxon-specific analyses of African inventory and other data<sup>12</sup> suggest that widespread changes in resource availability, such as increasing atmospheric carbon dioxide concentrations, may be the cause of the increase in carbon stocks<sup>13</sup>, as some theory<sup>14</sup> and models<sup>2,10,15</sup> predict.

Tropical forests cover 7–10% of the global land area, store 40–50% of carbon in terrestrial vegetation and annually process approximately six times as much carbon via photosynthesis and respiration as humans emit from fossil fuel use<sup>3,4</sup>. Tropical forests are therefore a

critical, yet very poorly quantified, component of the global carbon cycle. Although atmospheric CO<sub>2</sub> concentration data, combined with estimates of fossil fuel emissions, ocean carbon fluxes and carbon released from land-use change, indicate a global land carbon sink averaging 2.2 Pg C annually from 1980 to 2000, the location(s) of the sink has remained elusive<sup>1</sup>. Attempts to refine our understanding of these land sinks have predominantly focused on the interpretation of patterns of atmospheric CO<sub>2</sub> data using atmospheric transport models. However, studies remain poorly constrained and contradictory, largely because of a lack of atmospheric observations in the tropics<sup>15–17</sup>. A complementary approach towards solving the global carbon budget is to directly monitor specific land carbon pools. Over recent decades, long-term monitoring plots across Amazonia show that remaining old-growth forest trees are a sink of  $0.62 \pm 0.23 \text{ Mg Ch}^{-1} \text{ yr}^{-1}$  (0.5–0.8 Pg Cyr<sup>-1</sup> for all Amazonia<sup>6,7</sup>). However, it is not known if this is a regional or pan-tropical phenomenon because data from the largest tropical continent, Africa, have been almost completely absent<sup>9,10</sup>.

To assess whether similar changes are currently occurring in African tropical forest, we assembled existing long-term monitoring data and complementary field campaigns to create a ten-country network of monitoring plots called Afritron (African Tropical Rainforest Observation Network, [www.afritron.org](http://www.afritron.org)). Here we report data from 79 permanent sample plots spanning 40 years (1968–2007), located in closed-canopy moist forest, spanning West, Central and Eastern Africa (Supplementary Fig. 1, Supplementary Table 1). Our approach is to measure the diameter of all trees that are  $\geq 100$  mm in diameter in a given area (mean 2.1 ha) at least twice (mean interval 9.4 years), and apply allometric equations to convert tree-diameter measurements to carbon content, using standardized procedures (Methods). Following an empirical method to weight plot results relative to sampling effort (Methods, Supplementary Figs 4–10), we estimate that above-ground carbon stored in live trees averages  $202 \text{ Mg ha}^{-1}$  ( $n = 79$ ; bootstrapped

<sup>1</sup>Earth and Biosphere Institute, School of Geography, University of Leeds, Leeds LS2 9JT, UK. <sup>2</sup>Plant Systematic and Ecology Laboratory, University of Yaounde I, PO Box 047, Yaounde, Cameroon. <sup>3</sup>Resource Management Support Centre, Forestry Commission of Ghana, PO Box 1457, Kumasi, Ghana. <sup>4</sup>Department of Environmental Management and Toxicology, University of Agriculture, PMB 2240, Abeokuta, Ogun State, Nigeria. <sup>5</sup>Bureau Waardenburg bv, PO Box 365, 4100 AJ Culemborg, The Netherlands. <sup>6</sup>Institut de Recherche en Ecologie Tropicale (IRET), BP 7847, Libreville, Gabon. <sup>7</sup>SI/MAB Biodiversity Program, Smithsonian Institution, Suite 3123, 1100 Jefferson Drive SW, Washington DC 20560, USA. <sup>8</sup>Forest Ecology & Management Group, Department of Environmental Sciences, Wageningen University, PO Box 342, NL-6700, The Netherlands. <sup>9</sup>Plantlife International, 14 Rolleston Street, Salisbury, Wiltshire SP1 1DX, UK. <sup>10</sup>Wildlife Conservation Society-DRC, 1725 Avenue Monjibra, Chanic Building 2nd floor, Ngalinema, BP 240, Kinshasa I, Democratic Republic of Congo. <sup>11</sup>Eco-anthropologie et ethnobiologie, Département Hommes, Natures, Sociétés, MNHN, 4 av. du Petit Château, 91 800 Brunoy, France. <sup>12</sup>Centre for Ecology, Law and Policy, Environment Department, University of York, York YO1 5DD, UK. <sup>13</sup>Environmental Change Institute, School of Geography and the Environment, Oxford University, Oxford OX1 3QY, UK. <sup>14</sup>Department of Botany, University of Dar es Salaam, PO Box 35060, Dar es Salaam, Tanzania. <sup>15</sup>Centre for International Forestry Research, PO Box 0113 BOCBD, Bogor 16000, Indonesia. <sup>16</sup>Department of Plant & Soil Science, Cruickshank Building, School of Biological Sciences, University of Aberdeen, St Machar Drive, Aberdeen AB24 3UU, UK. <sup>17</sup>Department of Geography, Museum Building, School of Natural Sciences, Trinity College, University of Dublin, Dublin 2, Republic of Ireland. <sup>18</sup>Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, Ontario M5S 3B3, Canada. <sup>19</sup>Sommersbergseestr. 291, A-8990 Bad Aussee, Austria. <sup>†</sup>Present addresses: Inventory & Monitoring Program, National Park Service, 120 Chatham Lane, Fredericksburg, Virginia 22405, USA (J.A.C.); Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, PO Box 44, Kabale, Uganda (D.S.); Centre for International Forestry Research, PO Box 6596 JKPWB, Jakarta 10065, Indonesia (T.S.).

95% CI, 174–244; Supplementary Table 2). Carbon storage increased by  $0.63 \text{ Mg ha}^{-1} \text{ yr}^{-1}$ , with 73% of plots showing increased storage ( $n = 79$ ; bootstrapped 95% CI, 0.22–0.94; Fig. 1; 0.29% increase, CI, 0.05–0.50; Supplementary Table 2, see Supplementary Fig. 2 and Supplementary Table 2 for unweighted data). The distribution is non-normal, being a left-skewed leptokurtic distribution, because occasional disturbance events may sharply reduce carbon stocks in relatively small plots monitored over relatively short time periods (Fig. 1). Grouping plots into spatial clusters (Supplementary Fig. 1, Supplementary Table 2) gives similar results (Supplementary Information).

We monitored 163 ha for a combined census length of 744 years, but a simple scaling of the weighted mean increase in carbon stocks by total forest area may overestimate the actual carbon sink, because tree growth is relatively constant, yet mortality is highly variable and large mortality events are rare. Specifically, if there was no carbon sink and we sampled all African forest over a long enough period, the mean and median change in carbon storage would, by definition, be zero. However, when sampling over shorter time periods, most forest plots will be increasing in carbon storage (approximately constant growth), offset by occasional larger losses (stochastic mortality), thus we are sampling from a left-skewed long-tail distribution<sup>18</sup>. If the disturbance regimes in tropical forests are dominated by large and rare events then the tail of the distribution of change in carbon stocks would be very heavy, leading to likely positive biases in estimating changes in carbon stocks unless sample sizes are very large<sup>19</sup>. The only data on very large-scale mortality events come from a survey of Brazilian Amazonia ( $3.9 \times 10^6 \text{ km}^2$ ), which found only 19 large ( $>0.3 \text{ km}^2$ ) natural canopy gaps  $<2$  years old, totalling 0.0001% of the area surveyed, suggesting that large disturbances are very rare<sup>20</sup>. Furthermore, quantitative analyses of the statistical distributions of the large-scale Amazon data, less spatially extensive African data and carbon losses from mortality from this study are all consistent: mortality events in non-cyclone-affected tropical forests are dominated by small and frequent events (Supplementary Information and Supplementary Table 4). Forest simulations using these new analyses show that our estimates of changes in above-ground carbon stocks are unlikely to be biased<sup>19</sup>.

Yet, is our data set significantly different from that expected from a domain that is not increasing in carbon stocks? The ultimate cause of most large mortality events is the climate system producing extreme windstorms, rainfall events and droughts, often modelled using one of a family of long-tail generalized extreme value (GEV) distributions<sup>18</sup>. GEV theory<sup>18</sup> indicates that our data fit a Weibull distribution (Fig. 1, Methods). Shifting this distribution left to fix the mean change in carbon storage to zero provides a parsimonious ‘null model’ with which to compare the actual data (Supplementary Fig. 3, Methods). Re-sampling from this zero-mean distribution

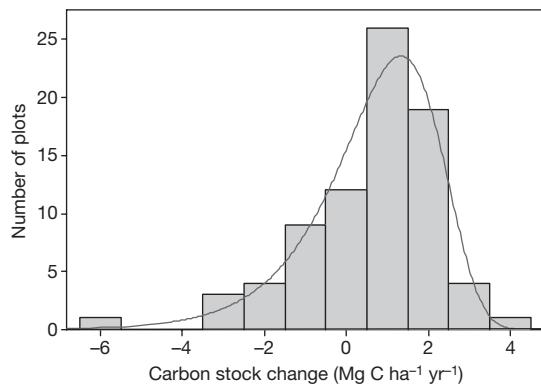
shows that it is highly unlikely that our data derive from a domain that was not increasing in carbon storage ( $P < 0.001$ ).

As the term ‘tropical forest’ does not have a universally agreed definition, because there are continua to sub-tropical, drier and open forest systems which are open to interpretation<sup>4</sup>, we estimate the total African closed-canopy forest sink using four widely used estimates of the extent of similar forest categories, giving an increase in above-ground live tree carbon stocks of  $0.24 \text{ Pg yr}^{-1}$  (CI, 0.08–0.35; Table 1). Assuming proportionate increases in (1) unmeasured below-ground live tree carbon stocks (because a larger mass of roots is required to support a larger mass of trees), (2) coarse woody inputs (necromass; because the mass of recently dead trees will be larger if the mass of live trees, of which a fraction die, is larger) and (3) stems smaller than our 100 mm diameter threshold, we estimate total tree carbon (live, dead, above- and below-ground) in African tropical forests to be increasing at a rate of  $0.34 \text{ Pg Cyr}^{-1}$  (CI, 0.15–0.43, Table 1). Of course, the limited number of sampling locations, absence of plots in some areas (notably central Democratic Republic of Congo, Supplementary Fig. 1) and differences in forest-area assessments all indicate that these first-estimate figures will require refinement in the future.

What are the implications of these findings for the carbon balance of Africa and the global carbon cycle more generally? The sink is of similar size to the estimated carbon flux from deforestation in tropical Africa<sup>21,22</sup> ( $0.1$ – $0.3 \text{ Pg yr}^{-1}$  over 1980–2000) and exceeds the flux from fossil fuel use in tropical Africa ( $0.04 \text{ Pg yr}^{-1}$  in 2000); also, it is consistent with a recent model prediction of a  $0.3 \text{ Pg Cyr}^{-1}$  sink in African forest vegetation<sup>10</sup>, and a recent African carbon cycle synthesis implying a terrestrial sink of  $0.4 \text{ Pg Cyr}^{-1}$  (ref. 9). More generally, combining all standardized inventory data from tropical Africa, America and Asia<sup>5,6,11</sup> (Methods) gives 156 inventory plots (562 ha monitored for a combined census length of 1,649 years), and a mean sampling effort-weighted increase in above-ground carbon stocks in live trees  $\geq 100$  mm of  $0.49 \text{ Mg Ch}^{-1} \text{ yr}^{-1}$  (95% CI, 0.29–0.66). This suggests a carbon sink over recent decades across all tropical forests of  $0.9 \text{ Pg Cyr}^{-1}$  (CI, 0.5–1.2), for above-ground live tree carbon, or  $1.3 \text{ Pg Cyr}^{-1}$  (CI, 0.8–1.6), for total tree carbon (Table 2, Supplementary Table 5). These estimates are consistent with a recent analysis of atmospheric CO<sub>2</sub> concentration data and atmospheric transport models that indicates a terrestrial tropical sink of  $\sim 1.5 \text{ Pg Cyr}^{-1}$  (ref. 15). Accounting for these results while balancing the global carbon cycle implies that the carbon source from land-use change, mostly tropical deforestation and degradation, is larger, and/or the carbon sink at temperate latitudes is smaller, than current central estimates suggest<sup>1</sup>.

The proximate cause of an Africa-wide increase in above-ground carbon storage could be a decrease in mortality rates, an increase in growth rates, or a combination of the two<sup>12</sup>. Long-term monitoring of sufficient plots will help distinguish amongst possibilities. The ultimate cause may be an increase in resource availability or, because trees are long-lived, these systems may still be changing *en masse* from past anthropogenic or natural impacts, or responding to a recent reduction in such disturbances<sup>23,24</sup>. Recent changes in resource availability include fertilization by anthropogenic additions of CO<sub>2</sub> to the atmosphere, changes in solar radiation at the Earth’s surface, increases in nutrient deposition rates and changes in rainfall<sup>13,25</sup>. Continental-scale disturbance-recovery cycles are less likely, because if Africa and Amazonia both experienced synchronous past disturbance events causing a carbon sink decades later, there would be indicative evidence, such as a large atmospheric CO<sub>2</sub> signal, of a past large-scale perturbation.

Nevertheless, anthropogenic pressures on African forests may have altered over time. For example, populations of large animals that disturb forests, such as elephants, have decreased, and increasing areas of forest have become protected, possibly also reducing disturbance levels. In addition, most African forest occurs close to the hydrological limits of closed-canopy forest, meaning that they are tightly coupled with climatic variability and change via hydrological



**Figure 1 | Histogram of annualized change in carbon stocks from 79 long-term monitoring plots across 10 countries in Africa.** Results presented are weighted by sampling effort (plot size and census-interval length), and fitted three-parameter Weibull distribution.

**Table 1 | Estimated carbon stocks and their annual increase for African tropical forest**

Study	Category	Area (10 <sup>6</sup> ha)	ALTC <sup>§</sup> (Pg)	TTC <sup>§</sup> (Pg)	ΔALTC ≥ 100 mm (Pg yr <sup>-1</sup> )	ΔALTC ≥ 10 mm (Pg yr <sup>-1</sup> )	ΔATC (Pg yr <sup>-1</sup> )	ΔTTC (Pg yr <sup>-1</sup> )
GLC2000	Humid tropical forest*	232.7	46.9 (40.5–56.8)	69.5 (60.9–80.7)	0.15 (0.05–0.22)	0.15 (0.06–0.23)	0.17 (0.08–0.25)	0.21 (0.09–0.27)
FRA CS	Closed forest†	352.7	71.1 (61.4–86.1)	105.3 (92.3–122.3)	0.22 (0.08–0.33)	0.23 (0.09–0.34)	0.26 (0.12–0.37)	0.32 (0.14–0.41)
FRA RS	Tropical forest	518.5	104.5 (90.2–126.5)	154.8 (135.6–179.8)	0.33 (0.11–0.49)	0.34 (0.13–0.50)	0.39 (0.17–0.55)	0.47 (0.21–0.60)
WCMC	Tropical forest‡	401.0	80.8 (69.8–97.8)	119.7 (104.9–139.1)	0.25 (0.08–0.38)	0.27 (0.10–0.39)	0.30 (0.13–0.42)	0.36 (0.16–0.46)
Mean		376.2	75.8 (65.5–91.8)	112.3 (98.4–130.5)	0.24 (0.08–0.35)	0.25 (0.09–0.37)	0.28 (0.12–0.40)	0.34 (0.15–0.43)

Estimates and 95% confidence intervals in parentheses are based on a measured stock of above-ground carbon in live trees  $\geq 100$  mm diameter of  $202 \text{ Mg C ha}^{-1}$  (CI, 174–244; for ALTC, TTC) and an increase of  $0.63 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$  (CI, 0.22–0.94; for ΔALTC, ATTC), and under the assumption that unmeasured minor vegetation components (small trees, shrubs, lianas), necromass and below-ground carbon in live trees increase proportionately (that is, are consistently assumed to be at equilibrium with the live trees  $\geq 100$  mm). Uncertainties are propagated as:  $\sqrt{(\text{error}_1)^2 + \dots + (\text{error}_n)^2}$ . GLC2000 (Global Land Cover Map 2000), FRA CS (Food and Agriculture Organisation Forest Resources Assessment 2000, Country Statistics), FRA RS (Food and Agriculture Organisation Forest Resources Assessment 2000, Remote Sensing statistics), all from ref. 8. WCMC, World Conservation Monitoring Centre (from [www.unep-wcmc.org](http://www.unep-wcmc.org)). ALTC, above-ground live tree carbon storage. ATC, above-ground tree carbon storage, including coarse necromass. TTC, total tree carbon storage, including below-ground carbon from tree roots. Trees  $\leq 100$  mm are 0.0519% (95% CI, 0.0166) of ALTC, from two African inventory plots<sup>11</sup>; coarse necromass is 0.127% (95% CI, 0.0319) of ALTC, from pan-Amazon forest average<sup>7</sup>; roots are 0.25% (95% CI, 0.10) of ALTC, from an African literature estimate<sup>30</sup>.

\* Also  $415.1 \times 10^6$  ha 'dry forest'.

† Also  $288.9 \times 10^6$  ha 'open forest'.

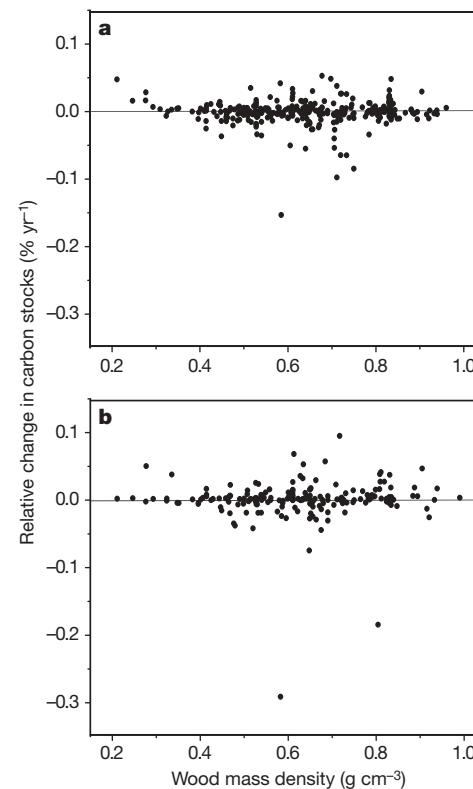
‡ Lowland evergreen, semi-evergreen, deciduous/semi-evergreen, and lower montane categories combined.

§ Scaled stock estimates from intact forest plots do not account for forest degradation and therefore will be modest overestimates.

changes. Today's African wet tropics have previously been climatically modestly drier at times over the past two to three millennia, which led to significant changes in vegetation, with much more open, disturbed, fire-prone, and less carbon-dense forests then covering many areas that are currently closed-canopy forest<sup>23,24</sup>. African forests are clearly sensitive to climatic changes, and past events may have left their signature on present day forests.

Taxon-level analyses can assist in distinguishing between the resource availability and disturbance-recovery hypotheses. If the increase in carbon stocks was caused by recovery from past disturbance events, then lighter-wooded species (lower wood mass density) would decline as a proportion of the forest stand, and heavier-wooded species increase, as is well known from ecological theory and observational evidence<sup>26</sup>. We would therefore predict a positive relationship between the change in the carbon stocks of a given taxon, relative to the forest stand, and its wood density. Alternatively, if increases in resource levels are the cause of the changes, then we predict either no relationship between the change in the carbon stocks of a taxon relative to the forest stand, and wood density, if all species benefit, or possibly an increase in lighter-wooded species relative to heavier-wooded species, because lighter-wooded species may dominate some types of more resource-rich environments<sup>27</sup>. Figure 2 shows that there is no relationship between the performance of taxa, relative to the stand as a whole, and their wood density; this suggests that the increase in mean carbon stocks across the 79 African plots has been caused, at least in part, by an increase in resource availability favouring species with a wide range of ecological habits, rather than recovery from past disturbance.

African tropical forests are providing important ecosystem services by storing carbon and being a carbon sink, thereby reducing the rate of increase of atmospheric CO<sub>2</sub>. With adequate protection these forests are likely to remain large carbon stores in the longer term. Securing this service will probably require formalizing and enforcing land rights for forest dwellers, alongside payments for ecosystem services to those



**Figure 2 | Relative change in carbon stocks and corresponding wood mass density values. a**, All 612 fully identified species occurring in  $\geq 5$  plots; **b**, all 200 fully identified genera occurring in  $\geq 5$  plots.

**Table 2 | Estimated annual increase in tropical forest carbon stocks**

Continent	Area (10 <sup>6</sup> ha)	ΔALTC $\geq 100$ mm (Pg yr <sup>-1</sup> )	ΔALTC $\geq 10$ mm (Pg yr <sup>-1</sup> )	ΔATC (Pg yr <sup>-1</sup> )	ΔTTC (Pg yr <sup>-1</sup> )
Central and South America	786.8	0.39 (0.23–0.52)	0.42 (0.27–0.56)	0.47 (0.31–0.61)	0.62 (0.39–0.73)
Africa	632.3	0.31 (0.18–0.42)	0.33 (0.20–0.43)	0.37 (0.24–0.48)	0.44 (0.26–0.53)
Asia	358.3	0.18 (0.10–0.24)	0.18 (0.11–0.25)	0.21 (0.13–0.27)	0.25 (0.15–0.30)
Total	1777.3	0.87 (0.52–1.17)	0.93 (0.58–1.24)	1.04 (0.68–1.35)	1.31 (0.79–1.56)

Estimates are based on a measured increase of above-ground carbon in live trees  $\geq 100$  mm diameter of  $0.49 \text{ Mg C ha}^{-1} \text{ yr}^{-1}$  (CI, 0.29–0.66). Areas are averages of 'tropical forest' (including more open and drier forest types) from four sources (GLC2000, FRA CS, FRA RS, WCMC). Full details of all forest types and sources of forest area data are in Supplementary Table 5. Abbreviations as in Table 1.

living near forested areas. Whether remaining intact forests will continue to sequester carbon, become neutral, or become a net source of carbon in the future is highly uncertain<sup>2,28</sup>. Improved monitoring and modelling of the tropical environment is required to better understand this trajectory.

## METHODS SUMMARY

All plots were selected in apparently mature closed-canopy forest  $\geq 0.2$  ha with diameters of all free-standing woody stems  $\geq 100$  mm measured using standardized methods within a given area at least twice and at least two years apart<sup>6</sup>, with 63 plots located at random within landscapes, and up to 16 plots placed in 'representative' areas of forest (see Supplementary Fig. 1 and Supplementary Table 1 for full details). Trees that increased in diameter  $>40$  mm  $\text{yr}^{-1}$  or shrunk 0.5 mm, including newly recruited stems  $>100$  mm in the final census, were assessed and errors corrected, as necessary (for example, interpolation between other censuses). For trees with changes in the point of measurement, we standardized growth rates obtained from the two points of measurement to a common line (Methods). We also calculated changes in carbon stocks using seven other methods to process trees with point of measurement changes (Supplementary Information). Diameter measurements were converted to carbon stocks using a published allometric equation for moist forests that includes terms for wood mass density and tree height<sup>29</sup>. We accounted for uncertainty associated with allometry using Monte Carlo techniques (Methods). Biomass was assumed to be 50% carbon<sup>29</sup>. We also calculated changes in carbon stocks using two other allometric equations (Supplementary Information). Seventy-nine random samples were taken from the zero-mean distribution 1,000 times, and the mean of each run calculated to compare with the actual mean from the data. Relative changes in species-level or genus-level C stocks were calculated as:  $[(\text{taxa\_C\_stock\_final} / \text{stand\_C\_stock\_final}) - (\text{taxa\_C\_stock\_initial} / \text{stand\_C\_stock\_initial})] / \text{census\_interval} \times 100$ . Only taxa identified with an accepted name in the African Flowering Plants database ([www.ville-ge.ch/cjb/bd/africa/](http://www.ville-ge.ch/cjb/bd/africa/)) were included in the analyses.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 9 June 2008; accepted 5 January 2009.

- Denman, K. L. et al. in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. et al.) 663–745 (Cambridge Univ. Press, 2007).
- Friedlingstein, P. et al. Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. *J. Clim.* **19**, 3337–3353 (2006).
- Malhi, Y. & Grace, J. Tropical forests and atmospheric carbon dioxide. *Trends Ecol. Evol.* **15**, 332–337 (2000).
- Lewis, S. L. Tropical forests and the changing Earth system. *Phil. Trans. R. Soc. Lond. B* **261**, 195–210 (2006).
- Phillips, O. L. et al. Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science* **282**, 439–442 (1998).
- Baker, T. R. et al. Increasing biomass in Amazonian forest plots. *Phil. Trans. R. Soc. Lond. B* **359**, 353–365 (2004).
- Phillips, O., Lewis, S. L., Baker, T. R., Chao, K.-J. & Higuchi, N. The changing Amazon forest. *Phil. Trans. R. Soc. B* **363**, 1819–1827 (2008).
- Mayaux, P. et al. Tropical forest cover change in the 1990s and options for future monitoring. *Phil. Trans. R. Soc. B* **360**, 373–384 (2005).
- Williams, C. et al. Africa and the global carbon cycle. *Carbon Balance Mgmt* **2**, 3, doi:10.1186/1750-0680-2-3 (2007).
- Ciais, P., Piao, S.-L., Cadule, P., Friedlingstein, P. & Chedin, A. Variability and recent trends in the African carbon balance. *Biogeosciences* **5**, 3497–3532 (2008).
- Chave, J. et al. Assessing evidence for a pervasive alteration of tropical tree communities. *PLoS Biol.* **6**, e45 (2008).
- Lewis, S. L. et al. Concerted changes in tropical forest structure and dynamics: Evidence from 50 South American long-term plots. *Phil. Trans. R. Soc. Lond. B* **359**, 421–436 (2004).
- Lewis, S. L., Malhi, Y. & Phillips, O. L. Fingerprinting the impacts of global change on tropical forests. *Phil. Trans. R. Soc. Lond. B* **359**, 437–462 (2004).
- Lloyd, J. & Farquhar, G. D. The CO<sub>2</sub> dependence of photosynthesis, plant growth responses to elevated atmospheric CO<sub>2</sub> concentrations and their interaction with soil nutrient status. 1. General principles and forest ecosystems. *Funct. Ecol.* **10**, 4–32 (1996).

- Stephens, B. B. et al. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO<sub>2</sub>. *Science* **316**, 1732–1735 (2007).
- Gurney, K. R. et al. Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models. *Nature* **415**, 626–630 (2002).
- Rodenbeck, C., Houweling, S., Gloo, M. & Heimann, M. CO<sub>2</sub> flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmos. Chem. Phys.* **3**, 1919–1964 (2003).
- Katz, R. W., Brush, G. S. & Parlange, M. B. Statistics of extremes: Modeling ecological disturbances. *Ecology* **86**, 1124–1134 (2005).
- Fisher, J. I., Hurtt, G. C., Thomas, R. Q. & Chambers, J. Q. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.* **11**, 554–563 (2008).
- Nelson, B. W. et al. Forest disturbance by large blowdowns in the Brazilian Amazon. *Ecology* **75**, 853–858 (1994).
- Houghton, R. A. & Hackler, J. L. Emissions of carbon from land use change in sub-Saharan Africa. *J. Geophys. Res.* **111**, G02003, doi:10.1029/2005JG000076 (2006).
- Achard, F., Eva, H. D., Mayaux, P., Stibig, H. J. & Belward, A. Improved estimates of net carbon emissions from land cover change in the tropics for the 1990s. *Glob. Biogeochem. Cycles* **18**, GB2008, doi:10.1029/2003GB002142 (2004).
- Brncic, T. M., Willis, K. J., Harris, D. J. & Washington, R. Culture or climate? The relative influences of past processes on the composition of the lowland Congo rainforest. *Phil. Trans. R. Soc. B* **362**, 229–242 (2007).
- Ngomanda, A. et al. Lowland rainforest response to hydrological changes during the last 1500 years in Gabon, Western Equatorial Africa. *Quat. Res.* **67**, 411–425 (2007).
- Wild, M. et al. From dimming to brightening: Decadal changes in solar radiation at Earth's surface. *Science* **308**, 847–850 (2005).
- Urquiza-Haas, T., Dolman, P. M. & Peres, C. A. Regional scale variation in forest structure and biomass in the Yucatan Peninsula, Mexico: Effects of forest disturbance. *For. Ecol. Mgmt* **247**, 80–90 (2007).
- Malhi, Y. et al. The above-ground coarse wood productivity of 104 Neotropical forest plots. *Glob. Change Biol.* **10**, 563–591 (2004).
- Cramer, W. et al. Tropical forests and the global carbon cycle: Impacts of atmospheric carbon dioxide, climate change and rate of deforestation. *Phil. Trans. R. Soc. Lond. B* **359**, 331–343 (2004).
- Chave, J. et al. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
- Deans, J. D., Moran, J. & Grace, J. Biomass relationships for tree species in regenerating semi-deciduous forest tropical moist forest in Cameroon. *For. Ecol. Mgmt* **88**, 215–225 (1996).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the villagers from Oban (Nigeria), Prince Town, Baneekurom and Dadioso (Ghana), Somolomo, Mekas, Lomie and Alat Makay (Cameroon) and Ekobakoba (Gabon) for hosting our fieldwork, and A. Moungazi, S. Mbadinga, H. Bourobou, L. N. Banak, T. Nzebi, K. Jeffery, SEGC/CIRMF/WCS Research Station Lopé, K. Ntim, K. Opoku, Forestry Commission of Ghana, T. Tafoek, Federal University of Agriculture, Abeokuta, Nigeria, ECOFAC-Cameroun, Cameroon Ministry Scientific Research and Innovation, Cameroon Ministry of Forests and Fauna (MINFOF), Forest Development Authority (Liberia), S. Grahame, R. Lowe, L. Banin and K. C. Nguembou for field assistance and logistical support. We thank R. Condit for helping access the Edoro and Lenda data sets, C. Chatelain for providing data from the African Flowering Plants database to standardize our species identifications, J. Chave for providing allometric data, L. Poorter for providing gap size-frequency data, B. Nelson for providing large-scale blow-down data, D. Appleyard for assistance with the plot map, and the Network for Vegetation Function (Nation Evolutionary Synthesis Center and ARC-NZ) for providing some wood mass density data. This work was supported by the Royal Society and the Natural Environment Research Council (UK).

**Author Contributions** S.L.L., O.L.P. and Y.M. conceived the experiment. S.L.L. designed the experiment. All authors except T.R.B., M.G., G.L-G., Y.M. and J.P. contributed materials. S.L.L., T.R.B., M.G., G.L-G., O.L.P. and J.P. provided analysis tools. S.L.L., G.L-G. and M.G. analysed data. S.L.L. wrote the paper. All authors commented on the analysis and presentation of the data.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.L.L. (s.l.lewis@leeds.ac.uk).

## METHODS

**Data collection and filtering.** Criteria for plot inclusion:  $\geq 0.2$  ha, clearly mixed-age stands (therefore apparently old-growth),  $\geq 50$  m from the anthropogenic forest edge, and free from major human impacts (for example, logged forest). Sixty-three plots were located at random with respect to gap-phase dynamics and up to 16 were located in ‘representative’ forest or avoided large tree-fall gaps (Supplementary Table 1). All free-standing woody stems  $\geq 100$  mm were measured, mapped and identified using standard forestry methods<sup>6</sup> for 71 plots. For eight plots published data were used (Supplementary Table 1). We identified 95%, 98% and 99% of stems to species, genus and family respectively. Trees that increased in diameter  $>40$  mm  $\text{yr}^{-1}$  or shrunk 0.5 mm were assessed to determine if they were inaccurately measured in the field (including new recruits). If judged inaccurate—for example, a slow-growing species growing abnormally fast—the diameter was interpolated or extrapolated using data from the same stem measured over other censuses, or, if only one accurate measurement was available, by substituting the median growth rate for the same diameter size-class from that plot (100–199, 200–399 and 400+ mm diameter). The median was chosen (1) to be conservative with respect to the hypothesis being tested, and (2) because sample sizes can be small for the largest size-class, hence the mean may not be robust. The standard method of tree measurement is ‘diameter at breast height’ (d.b.h.) defined as 1.3 m from the base of the stem, if approximately cylindrical at this point. If the transverse section of the stem is not cylindrical at 1.3 m height—owing to buttresses or other stem deformities—the measurement location is raised to a point where it is. Thus, as some trees grow, buttresses and other stem deformities may develop, and therefore occasionally the height of the point of measurement (POM) must be raised, so the measurement is consistent with standardized protocols. Tree growth estimates are therefore sometimes complicated by different POMs on the same tree, because changes in a POM can cause large negative growth in individuals that have actually grown (and hence needed a POM change). Our approach to trees with changing POMs was to utilize the greatest amount of information about each stems’ growth rate to estimate its size. We harmonize the two disjointed sets of growth measurements (from the original POM, and the new POM) by replacing the measured diameters with the mean of (1) the ratio of the original to the new POM diameter measurements (to standardize each diameter measurement to the height of the original POM), and (2) the ratio of the final to the original POM diameter measurement (to standardize each diameter measurement to the height of the final POM). For comparison, we plot seven other methods of dealing with POM changes, including the commonest method—of substituting mean growth rates from the same diameter size-class trees from the plot—which give almost identical results (Supplementary Information and Supplementary Fig. 12).

**Carbon stock estimation.** Diameter measurements were converted to carbon stocks using a published allometric equation for moist forests including terms for wood mass density and tree height, with biomass assumed to be 50% carbon<sup>29</sup>. A wood density database was compiled from published sources (737 species; Supplementary Table 3). Species in both the wood density and plot databases were standardized for orthography and synonymy using the African Flowering Plants Database ([www.ville-ge.ch/cjb/bd/africa/](http://www.ville-ge.ch/cjb/bd/africa/)) to maximize matches between the databases. Stems were matched to species-specific wood density values if

possible (62% of stems), or the mean values for the genera (23%), or family (11%), as appropriate. For incompletely identified stems, we used the next higher known taxonomic category or the plot-mean wood density (5% of stems). Tree heights were estimated from a height–diameter relationship derived from the measurements of 1,982 trees in African forests (Supplementary Fig. 14). Uncertainty associated with the allometric equation was propagated using Monte Carlo techniques. We used the residuals from the allometric equation to bootstrap the distribution of errors associated with allometry, and similarly for the height–diameter relationship. We then calculated the carbon in each tree with added random error from the bootstrapped distributions, summed over each plot for each census and the change in carbon stocks  $\text{ha}^{-1} \text{yr}^{-1}$  calculated (using Matlab). We repeated this process of adding random error to each tree and summing over each plot for each census 100 times, from which we calculated a mean change in carbon stocks for each plot from the 100 perturbed samples. Supplementary Table 2 reports initial, final and change in C stocks for the actual data and means of 100 perturbed samples. For five plots we used published data where only basal area was available (Supplementary Tables 1, 2). We converted these to carbon stocks by correlating basal area and carbon stocks for the 71 plots with full tree-by-tree data. For three published plots we used the published carbon stock values (Supplementary Tables 1, 2).

**Weighting analysis.** Optimum weightings of each replicate corresponding to the sampling effort employed (plot size and length of monitoring period) were derived empirically, assuming *a priori* that there is no pattern in the change in carbon stocks with monitoring period or plot size, by assessing patterns in the residuals of sampling effort versus carbon storage change, following different weightings. Weighting by the square root of the number of years of monitoring removes any pattern in the residuals. Similarly, weighting by the cube root of plot size removes any pattern in the residuals (Supplementary Figs 4–7). These results suggest that these square root and cube root relationships best describe how census length or plot size are related to sampling error, respectively. These results appear to be robust and general, as they are repeated when the Africa data are combined with other standardized data and re-analysed<sup>6,11</sup> (Supplementary Figs 7, 8). The final weighting of the square root of the number of years of monitoring plus cube root of plot size minus one (to avoid double-accounting) shows no pattern in the residuals (Supplementary Fig. 10). Weighting by sampling effort gives a similar mean but reduces the left-skew, because longer monitoring periods and larger plots dampen the impacts of occasional mortality events on underlying carbon storage trends (compare Fig. 1 with Supplementary Fig. 2).

**Generalized extreme value and re-sampling analysis.** We fitted a GEV distribution using the ExtRemes package<sup>18</sup> in R. It shows that the shape parameter from the distribution is significantly less than zero, indicating the data best fit a Weibull (stretched exponential) distribution<sup>18</sup> ( $-0.465$ ; 95% CI,  $-0.556$  to  $-0.371$ ). We then fitted the three-parameter Weibull probability density function to the data and shifted this distribution to have a mean change in carbon stocks of zero, our ‘null distribution’. From this distribution we re-sampled 79 values 1,000 times and calculated the mean change in carbon stocks for each run to compare with the mean from our data. The actual mean from our data was larger than any single value from our 1,000 re-samples from the null distribution.

## LETTERS

# Nodal signalling is involved in left-right asymmetry in snails

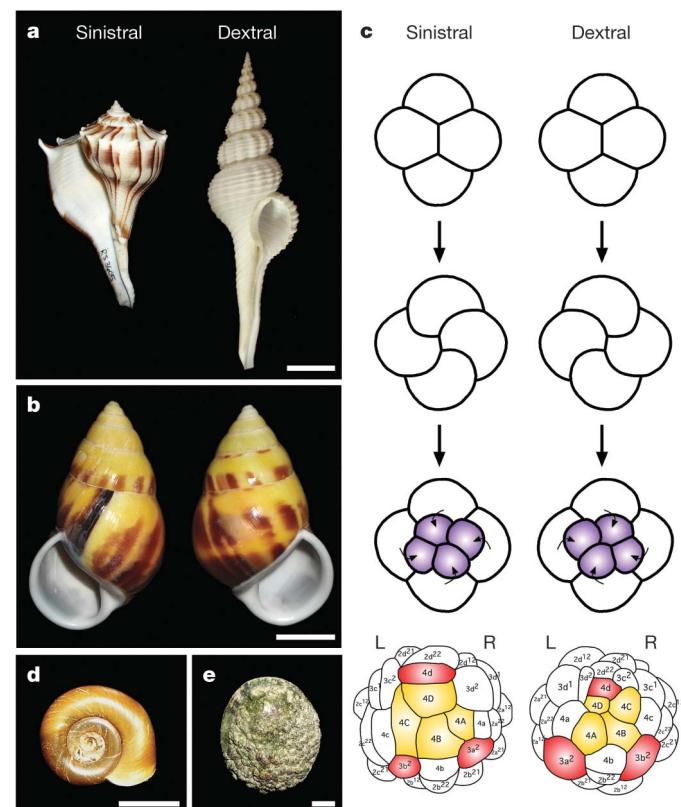
Cristina Grande<sup>1,2,3</sup> & Nipam H. Patel<sup>1,2,3</sup>

Many animals display specific internal or external features with left-right asymmetry. In vertebrates, the molecular pathway that leads to this asymmetry uses the signalling molecule Nodal, a member of the transforming growth factor- $\beta$  superfamily<sup>1</sup>, which is expressed in the left lateral plate mesoderm<sup>2</sup>, and loss of *nodal* function produces a randomization of the left-right asymmetry of visceral organs<sup>3,4</sup>. Orthologues of *nodal* have also been described in other deuterostomes, including ascidians and sea urchins<sup>5,6</sup>, but no *nodal* orthologue has been reported in the other two main clades of Bilateria: Ecdysozoa (including flies and nematodes) and Lophotrochozoa (including snails and annelids). Here we report the first evidence for a *nodal* orthologue in a non-deuterostome group. We isolated *nodal* and *Pitx* (one of the targets of Nodal signalling) in two species of snails and found that the side of the embryo that expresses *nodal* and *Pitx* is related to body chirality: both genes are expressed on the right side of the embryo in the dextral (right-handed) species *Lottia gigantea* and on the left side in the sinistral (left-handed) species *Biomphalaria glabrata*. We pharmacologically inhibited the Nodal pathway and found that *nodal* acts upstream of *Pitx*, and that some treated animals developed with a loss of shell chirality. These results indicate that the involvement of the Nodal pathway in left-right asymmetry might have been an ancestral feature of the Bilateria.

Asymmetric expression and action of the Nodal pathway is a conserved feature of deuterostomes, with *nodal* and its target gene *Pitx* acting on the left side of the embryo in chordates but on the right side in echinoderms (sea urchins)<sup>7</sup>. Until now, the lack of identified *nodal* orthologues in the genomes of cnidarians and ecdysozoans suggested that *nodal* evolved in an ancestor within the early deuterostome lineage. Furthermore, although the molecular pathways involved in several asymmetries in flies and nematodes are being actively investigated, so far the Nodal pathway does not seem to be involved. The third main group of Bilateria, the Lophotrochozoa, which includes snails and annelids, also displays morphological asymmetries. One of the asymmetries that has particularly intrigued researchers is snail chirality, which refers to the body handedness and direction of shell coiling (Fig. 1). Dextral (right-handed) and sinistral (left-handed) forms can be found at many taxonomic levels (Fig. 1a, b), although most extant snail species are dextral<sup>8</sup>. Body chirality can be related to a much earlier chirality that is seen in all gastropod embryos, namely the chirality of the spiral cleavage pattern seen at the transition from the 4- to 8-cell embryo (Fig. 1c). For some species, both chiral forms exist within a population (Fig. 1b), and body handedness is determined by a single locus that functions maternally to determine the chirality of offspring<sup>9–11</sup>.

To gain insight into the molecular pathway for left-right asymmetry in gastropods, we investigated some of the previously described genes critical for left-right determination in other organisms. Specifically, we focused on members of the Nodal pathway in two species of snails that differ in chirality: the sinistral species

*Biomphalaria glabrata* (Fig. 1d), an intermediate host for the pathogen that causes schistosomiasis, and the dextral species *Lottia gigantea* (Fig. 1e), the complete genome sequence of which has been assembled. We first found in *L. gigantea* several components of the Nodal signalling pathway, including *nodal* itself as well as one of the known target genes of *nodal*, the paired homeobox gene *Pitx*. We subsequently characterized complementary DNAs encoding proteins



**Figure 1 | Chirality in snails.** **a**, Species with different chirality: sinistral *Busycon pulleyi* (left) and dextral *Fusinus salisbury* (right). **b**, Sinistral (left) and dextral (right) shells of *Amphidromus perversus*, a species with chiral dimorphism. **c**, Early cleavage in dextral and sinistral species (based on ref. 27). In sinistral species, the third cleavage is in a counterclockwise direction, but is clockwise in dextral species. In the next divisions the four quadrants (A, B, C and D) are oriented as indicated. Cells coloured in yellow have an endodermal fate and those in red have an endomesodermal fate in *P. vulgata* (dextral)<sup>15</sup> and *B. glabrata* (sinistral)<sup>28</sup>. L and R indicate left and right sides, respectively. **d**, *B. glabrata* possesses a sinistral shell and sinistral cleavage and internal organ organization. **e**, *L. gigantea* displays a dextral cleavage pattern and internal organ organization, and a relatively flat shell characteristic of limpets. Scale bars: **a**, 2.0 cm; **b**, 1.0 cm; **d**, 0.5 cm; **e**, 1.0 cm.

<sup>1</sup>Department of Molecular and Cell Biology, <sup>2</sup>Department of Integrative Biology, and <sup>3</sup>Center for Integrative Genomics, University of California, Berkeley, California 94720-3200, USA.

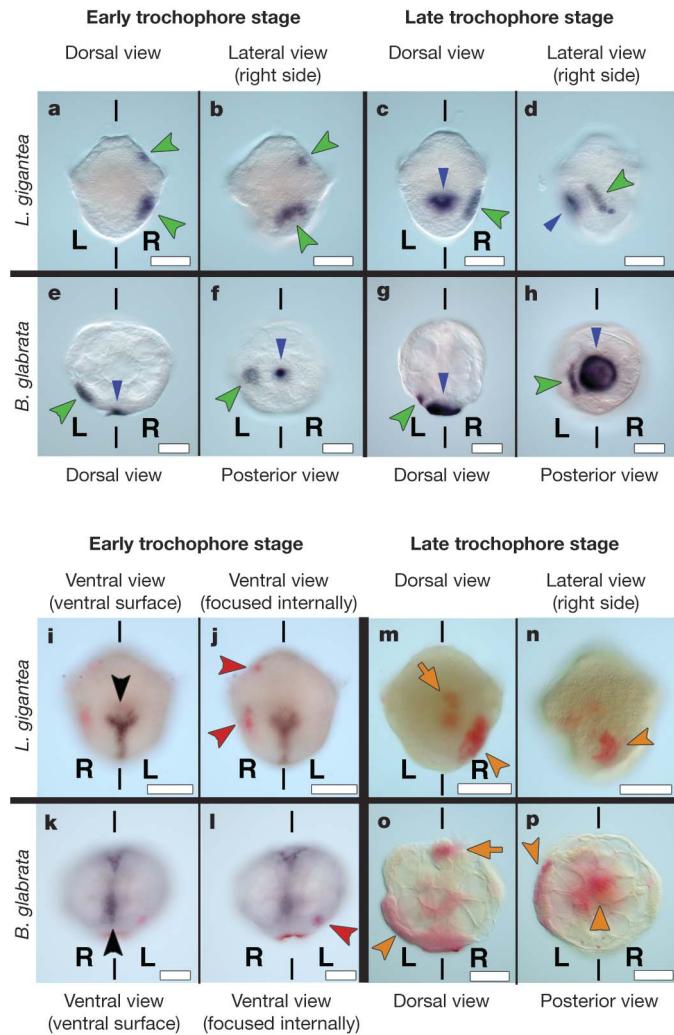
related to Nodal and Pitx in both *L. gigantea* and *B. glabrata*. Bayesian phylogenetic analyses confirmed, with high statistical support, that these potential orthologues of Nodal are indeed more closely related to deuterostome Nodal than to any other transforming growth factor- $\beta$  (TGF- $\beta$ ) family member (Supplementary Figs 1a and 2). To determine whether *nodal* is also present in other lophotrochozoans, we also identified a potential orthologue of *nodal* in the annelid *Capitella* species I, and further phylogenetic analyses confirmed the close relationship between this sequence and those of the gastropod and deuterostome Nodal proteins (Supplementary Figs 1a and 2). In addition, a single potential orthologue of *Pitx* was identified in both snail species (Supplementary Fig. 1b).

We then examined the expression pattern of *nodal* and *Pitx* in both *L. gigantea* and *B. glabrata*. Snail embryos undergo indirect development, producing first a trochophore and then a veliger larva. We first examined the trochophore stage because the anterior-posterior and dorsal-ventral axes at this time are morphologically clear, and found that both genes were expressed in left-right asymmetric patterns (Fig. 2). In *L. gigantea* there are two ectodermal domains of *nodal* expression on the right side of the embryo: one cephalic region plus a lateral domain near where shell formation initiates (Fig. 2a-d). In *B. glabrata*, expression is seen in an ectodermal region on the left side of the embryo near where the shell gland will form in the posterior midline region (Fig. 2e-h). In the veliger stage of *B. glabrata*, an additional domain in the left cephalic ectoderm is seen (Supplementary Fig. 3). To confirm the right versus left expression, we double-labelled embryos of both species for *nodal* and hedgehog, a gene that is expressed at the ventral midline in snails (Fig. 2i-l) <sup>12</sup>.

At the trochophore stage of *L. gigantea*, *Pitx* expression is seen in a group of ectodermal cells on the right side of the larvae, adjacent to those that express *nodal*, as well as in the developing gut (Figs 2m, n and 3i, j). At later stages, *Pitx* expression is maintained in these domains but also becomes visible in the right cephalic ectoderm and in a symmetric domain of the visceral mass (Supplementary Fig. 4). In the trochophore of *B. glabrata*, *Pitx* expression is observed in the ectoderm on the left side of the larvae close to the shell gland and adjacent to the group of cells that express *nodal*, as well as in symmetric patterns in the stomodeum and gut (Fig. 2o-p). Later in development, *Pitx* is also symmetrically expressed in the developing cephalic tentacles (Supplementary Fig. 4).

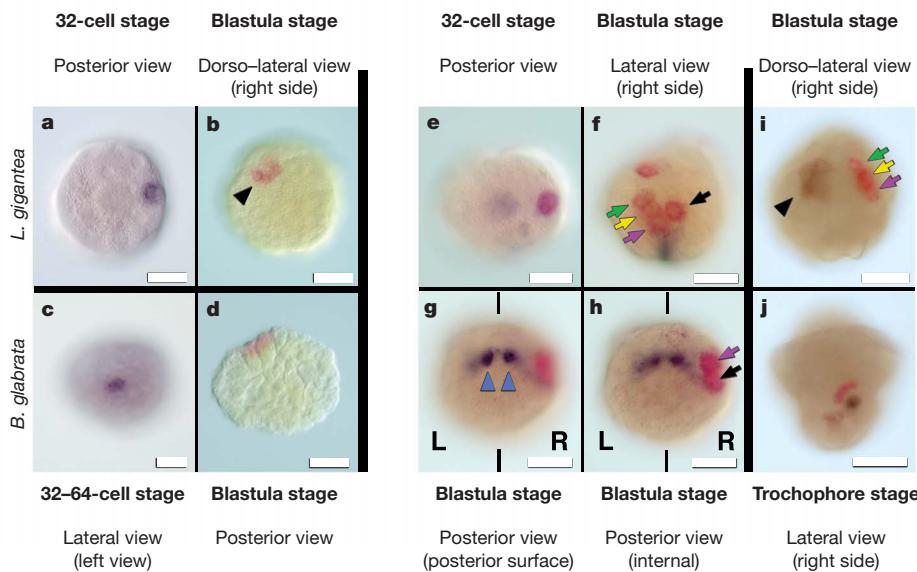
We next examined earlier stages of development to determine when the asymmetric patterns were first visible. In both species, *nodal* transcripts were first detected at the 32-cell stage (Fig. 3a, c). The bilateral axis is visible in snails at the 32- to 64-cell stage, when the macromere of the D quadrant, 3D, produces mesentoblast 4d, which subsequently divides in a bilateral fashion to form paired stem cells that will give rise to the mesodermal germ bands <sup>13</sup>. Embryos of *L. gigantea* double-labelled for *nodal* and *brachyury*, a gene that is expressed in 3D, mesentoblast 4d and then very strongly in the distinctly left-right bilateral cells 3d<sup>2</sup> and 3c<sup>2</sup> (ref. 14), show that the expression of *nodal* is clearly left-right asymmetric at the 32–64-cell stage and beyond (Fig. 3e-h). Using the *P. vulgata* fate map <sup>15</sup> as a guide, we conclude that *nodal* is expressed in the C quadrant, specifically in the derivatives of the micromeres 1c and 2c. In *P. vulgata*, the progeny of 1c are part of the right cephalic ectoderm of the larvae, whereas 2c-derived cells are part of the ectoderm of the right side of the foot, the mantle fold and the shell field <sup>15</sup>, suggesting that the two ectodermal domains seen at the larval stages are composed of a subset of the nodal-expressing cells from the 64-cell stage. *Pitx* expression was first detected in both snail species at the 64-cell stage in a group of cells of the D quadrant, close to those that express *nodal* (Fig. 3b, d, i and j).

To investigate the function of Nodal signalling in snail development experimentally, we used the chemical inhibitor SB-431542. The TGF- $\beta$  superfamily includes a large number of ligands, but, because the drug SB-431542 specifically interferes with type I receptors Activin receptor-like kinase 4 (Alk4), Alk5 and Alk7 (ref. 16), only the activity of



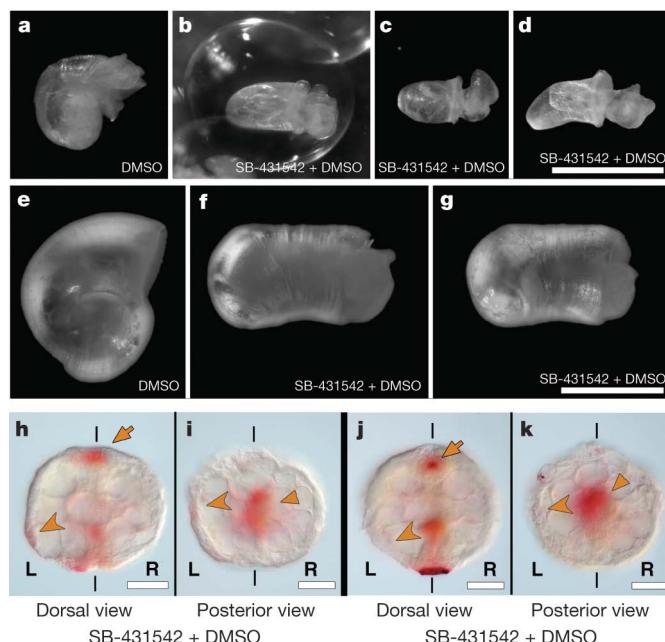
**Figure 2 | *nodal* and *Pitx* expression in snails.** Anterior is up, L and R indicate left and right sides. The blue arrowhead in c-h indicates non-specific staining of the shell. **a, b**, *nodal* is expressed in the right cephalic region (upper green arrowhead) and in the right lateral ectoderm (lower green arrowhead) in *L. gigantea*, as seen from dorsal (**a**) and right lateral (**b**) views. **c**, Expression is maintained in the right lateral ectoderm (green arrowhead); the right lateral view (**d**) shows that *nodal* expression (green arrowhead) is near the right side of the developing shell (blue arrowhead). **e-h**, *nodal* is expressed in the left lateral ectoderm (green arrowhead) in *B. glabrata*, as seen from dorsal (**e**) and posterior (**f**) views; **g, h**, Expression is maintained in the left lateral ectoderm (green arrowhead); the posterior view (**h**) shows that *nodal* expression (green arrowhead) is near the left side of the developing shell (blue arrowhead). **i-l**, *hedgehog* (black arrowheads in **i** and **k**) is expressed along the ventral midline, and *nodal* (red arrowheads in **j** and **l**) is expressed on the right side of *L. gigantea* (**j**) and on the left side of *B. glabrata* (**l**). **m, n**, *Pitx* is expressed in the visceral mass (orange arrow) and right lateral ectoderm (orange arrowhead) in *L. gigantea*, as seen from dorsal (**m**) and right lateral (**n**) views. **o, p**, *Pitx* is expressed in the stomodeum (orange arrow), visceral mass (orange triangle) and the left lateral ectoderm (orange arrowhead) in *B. glabrata*, as seen from dorsal (**o**) and posterior (**p**) views. Scale bars: 50  $\mu$ m.

Nodal, Activin and TGF- $\beta$  subfamily ligands are expected to be blocked by this drug. Of the potential members of the TGF- $\beta$  superfamily that we have identified in the genome of *L. gigantea* (Supplementary Results and Supplementary Fig. 2), only Nodal and Activin signalling should be affected by SB-431542. Given, however, that we have not been able to detect the expression of any of the potential *activin* homologues during embryonic or larval stages, we suggest that SB-431542 treatment might affect Nodal signalling exclusively during snail development before the juvenile stage.



**Figure 3 | Early expression of *nodal* and *Pitx* in snails.** **a**, 32-cell stage *L. gigantea* expressing *nodal* in a single cell. **b**, Group of cells expressing *Pitx* in *L. gigantea*. **c**, Onset of *nodal* expression in *B. glabrata*. **d**, A group of cells expressing *Pitx* in *B. glabrata*. **e**, 32-cell *L. gigantea* expressing *nodal* (red) in a single cell (2c) and *brachyury* (black) in two cells (3D and 3c). **f–h**, *brachyury* (black) is expressed in a symmetrical manner in progeny of 3c and 3d blastomeres (blue triangles in g), thus marking the bilateral axis, and *nodal* (red) is expressed on the right side of *L. gigantea* in the progeny of 2c

Different developmental stages of *B. glabrata* were treated with SB-431542, and the percentages of specific abnormalities varied depending on the concentration and the timing of drug treatment (Supplementary Table 1). The most frequently observed phenotype after application of drug at early stages (1–16 cells) was the production of embryos that failed to complete gastrulation and thus could not form juveniles with shells. Some of the embryos that successfully completed gastrulation, however, displayed a striking phenotype—they developed with non-coiled shells (Fig. 4a–g). At 5 μM drug concentration, the straight-shelled phenotype was seen in 8% of the animals that managed to complete gastrulation, but at 10 μM concentration this value rose to 43% of those that completed gastrulation.



and 1c blastomeres, as seen from the lateral (f) and posterior (g, h) views of the same embryo. **i**, A group of cells expressing *nodal* (red) in the C quadrant and *Pitx* (black) in the D quadrant of the 120-cell-stage embryo of *L. gigantea*. **j**, *nodal* (red) and *Pitx* (black) expression in adjacent areas of the right lateral ectoderm in *L. gigantea*. L and R indicate the left and right sides of the embryo, respectively. The black triangle in b and i, the green, yellow and pink arrows in f and i, and the black and pink arrows in f and h point to the equivalent cells. Scale bars: 50 μm.

Although the shells showed a slight dorsal–ventral curvature, they were tubular with no sign of the left-handed coiling usually seen in *B. glabrata* (Fig. 4a–g). Varying the time of drug application revealed that exposure at the trophophore stage or later yielded animals that were almost always normal, and exposure at the blastula or gastrula stage yielded no coiling defects. Only by applying the drug before the blastula stage could we obtain animals with coiling defects, suggesting that a role in left–right asymmetry might be the earliest function of *nodal* within the embryo (see Supplementary Table 1). Efforts to further narrow the time of drug inhibition by washing off the drug were ineffective, possibly owing to the slow diffusion rate of the drug through the egg capsule or a slow reversal of the biochemical effect of the drug on the receptors.

We also followed the development of some of the animals with straight shells over the course of several days. We found that these animals continued to enlarge their shells, indeed forming shells that were quite long and robust, but remained straight (Supplementary Fig. 5). To rule out the possibility that the lack of shell coiling was simply due to general poor growth, we applied a different chemical

**Figure 4 | Wild-type coiled and drug-treated non-coiled shells of *B. glabrata* and *Pitx* expression in drug-treated embryos.** Control animals (**a**, **e**) display the normal sinistral shell morphology. Drug-treated animals (**b–d**, **f–g**, exposed to SB-431542 from the 2-cell stage onwards) have straight shells. **b–d** show three different living individuals; **f** and **g** are a fourth individual, ethanol-fixed, and shown from the side (**f**) and slightly rotated (**g**). **h–k**, *Pitx* expression in embryos exposed to SB-431542. Dorsal (**h**) and posterior (**i**) views of an embryo showing reduced levels of expression. *Pitx* expression is maintained in the stomodeum (orange arrow in **h**) and the visceral domain (orange triangle in **i**), but asymmetric expression in left ectoderm is greatly reduced (orange arrowhead). Dorsal (**j**) and posterior (**k**) views of an embryo in which the asymmetric ectodermal expression of *Pitx* is undetectable (orange arrowhead in **j** and **k** show where expression would be expected), although the stomodeal (orange arrow in **j**) and visceral (orange triangle in **k**) domain expression of *Pitx* is normal. *Pitx* expression levels shown in **h–k** should be compared to levels in wild-type embryos in Fig. 2o and p, which are the same levels seen in DMSO-treated animals. L and R indicate the left and right sides of the embryo, respectively. Scale bars: **a–d**, 1.0 mm; **e–g**, 0.5 mm; **h–k**, 50 μm.

inhibitor, rapamycin—a drug that interferes with the cellular metabolic machinery modifying cell growth and proliferation (see Methods Summary). Although the treated embryos showed various, sometimes severe, defects in their morphology, the shells that were formed always showed some degree of coiling. Non-coiled shells similar to those recovered with the SB-431542 drug treatment were not detected (Supplementary Table 2).

To test the effects of the SB-431542 drug treatment on the Nodal pathway at the molecular level, we compared the levels of expression of *nodal* and *Pitx* in control and drug-treated embryos. Although no variation in levels of *nodal* expression was detected (both control and drug-treated embryos that survived to the trophophore stage showed strong *nodal* expression), in 30% of the treated embryos the level of expression in the asymmetric domain of *Pitx* was greatly reduced (Fig. 4h, i and Supplementary Table 3) and in 9.5% of the treated embryos asymmetric expression was undetectable (Fig. 4j, k and Supplementary Table 3). Only the asymmetric expression of *Pitx* was affected by the drug treatment; expression levels in the symmetric domains in the stomodeum and gut were unaltered, suggesting separate regulatory elements for these domains of expression, as previously described in deuterostomes<sup>17</sup>. These results indicate that *Pitx* is downstream from *nodal* in snails, as in deuterostomes. However in snails, unlike in deuterostomes, *nodal* does not seem to regulate its own expression. This is consistent with our observation that *nodal* expression in snails is asymmetric from the outset, whereas in vertebrates *nodal* expression is initially symmetric and depends on the regulated feedback of *nodal* signalling to achieve an asymmetric pattern<sup>2</sup>.

The drug treatment results, together with our analysis of the expression of *nodal* and *Pitx*, provide preliminary support for our contention that Nodal signalling has a role in left-right asymmetry in snails. The reduction of Nodal signalling leads to a randomization of asymmetry in vertebrates, but it is interesting to note that we observe a lack of asymmetry in snails. Although chirality in snails is first defined at the transition from the 4- to the 8-cell stage, the first indication of morphological asymmetry in snails is given by a displacement of the shell gland to the left in dextral species and to the right in sinistral species. We suggest that the asymmetric activity of the Nodal pathway could lead to unequal formation of shell-producing cells on the two sides of the embryo, or could asymmetrically alter the rate of shell production.

Our results provide new insights into the evolution of body plans and left-right specification in Bilateria. Previous studies suggest that some general mechanisms (at the level of involvement of gap junctions and ( $H^+ + K^+$ )ATPase activity) contributing to left-right asymmetry are shared between distant phyla<sup>18,19</sup>. We hypothesize that there is an even closer linkage. Although *Pitx* orthologues have also been identified in non-deuterostomes such as *Drosophila melanogaster* and *Caenorhabditis elegans*, in these species *Pitx* has not been reported in asymmetrical expression patterns. Our results suggest that asymmetrical expression of *Pitx* might be an ancestral feature of the bilaterians. Furthermore, our data suggest that *nodal* was present in the common ancestor of all bilaterians and that it too may have been expressed asymmetrically. Various lines of evidence indicate that the last common ancestor of all snails had a dextral body<sup>20</sup>. If this is true, then our data would suggest that this animal expressed both *nodal* and *Pitx* on the right side. Combined with the fact that *nodal* and *Pitx* are also expressed on the right side in sea urchins<sup>7,21</sup>, this raises the possibility that the bilaterian ancestor had left-right asymmetry controlled by *nodal* and *Pitx* expressed on the right side of the body. Although independent co-option is always a possibility, the hypotheses we present can be tested by examining *nodal* and *Pitx* expression and function in a variety of additional invertebrates.

Our data also provide a molecular entry point into understanding chirality in snails. In vertebrates, the actual symmetry-breaking event occurs before *nodal* is asymmetrically expressed, and the mechanisms that break symmetry seem to be different between various vertebrate species<sup>2,22–24</sup>. Likewise, in snails symmetry must be broken before the

8-cell stage, before *nodal* expression begins. Chirality in snails is determined by a still uncharacterized maternal factor, but once chirality is established *nodal* and *Pitx* are expressed on one side of the embryo. Future studies that determine how *nodal* expression is regulated will lead us towards an understanding of the actual symmetry-breaking event in snails, and examination of the steps downstream of Nodal signalling will provide new insights into the developmental control of complex animal morphologies such as shell coiling.

## METHODS SUMMARY

We searched the National Center for Biotechnology Information (NCBI) database of genomic sequences from *L. gigantea*, and found nine potential members of the TGF-β superfamily. Phylogenetic analyses included the newly determined amino acid sequences of Nodal from *L. gigantea*, *B. glabrata* and *Capitella* sp. I, as well as the Nodal sequences of other deuterostomes and other TGF-β superfamily members available from GenBank. Sequence data were analysed with MacClade version 4.05 OSX, Clustal X version 1.62b, and PAUP\* version 4.0b10. Gene orthology was determined by Bayesian-inference-based methods with MrBayes 3.12. Further details of phylogenetic analyses are available on request.

Sexually mature *L. gigantea* were collected in Los Angeles, California, during the breeding season, and *in vitro* fertilizations were performed as described previously<sup>25</sup>. A breeding population of *B. glabrata* is maintained in freshwater tanks at 25 °C, and embryos and larvae were regularly collected and raised in the laboratory. Embryos of both species were fixed as described previously<sup>26</sup>. PCR reactions were performed with gene-specific (for *L. gigantea*) and degenerate (for *B. glabrata*) primer sequences (available from the authors on request). 5' and 3' RACE PCR was performed with Invitrogen RACE reagents. *In situ* hybridizations were performed with digoxigenin- and fluorescein-labelled RNA probes as described previously<sup>26</sup>. Nodal inhibition was performed by placing egg masses of *B. glabrata* in fresh water containing 1% DMSO and SB-431542 (TOCRIS Bioscience) at concentrations of 5 μM or 10 μM (diluted from a 1 mM stock of SB-431542 in DMSO); treatment with rapamycin was performed by placing egg masses in fresh water containing 1% DMSO and rapamycin (Eton Bioscience Inc.) at 10 μM (diluted from a 10 mM stock of rapamycin in DMSO). Control embryos were exposed to 1% DMSO. All controls and drug-treated embryos were kept in the dark during the treatment period.

Received 29 August; accepted 31 October 2008.

Published online 21 December 2008.

- Massagué, J. & Gomis, R. R. The logic of TGFβ signaling. *FEBS Lett.* **580**, 2811–2820 (2006).
- Hamada, H., Meno, C., Watanabe, D. & Saijoh, Y. Establishment of vertebrate left-right asymmetry. *Nature Rev. Genet.* **2**, 103–113 (2002).
- Supp, D. M., Witte, D. P., Potter, S. S. & Brueckner, M. Mutation of an axonemal dynein affects left-right asymmetry in *inversus viscerum* mice. *Nature* **389**, 963–966 (1997).
- Okada, Y. et al. Abnormal nodal flow precedes *situs inversus* in iv and inv mice. *Mol. Cell* **4**, 459–468 (1999).
- Morokuma, J., Ueno, M., Kawanishi, H., Saiga, H. & Nishida, H. *HrNodal*, the ascidian nodal-related gene, is expressed in the left side of the epidermis, and lies upstream of *HrPitx*. *Dev. Genes Evol.* **212**, 439–446 (2002).
- Duboc, V., Rottinger, E., Besnard, L. & Lepage, T. *Nodal* and *BMP2/4* signaling organizes the oral-aboral axis of the sea urchin embryo. *Dev. Cell* **6**, 397–410 (2004).
- Duboc, V., Rottinger, E., Lapraz, F., Besnard, L. & Lepage, T. Left-right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. *Dev. Cell* **9**, 147–158 (2005).
- Schilthuizen, M. & Davison, A. The convoluted evolution of snail chirality. *Naturwissenschaften* **92**, 504–515 (2005).
- Boycott, A. E. & Diver, C. On the inheritance of sinistrality in *Limnaea peregra*. *Proc. R. Soc. Lond. B* **95**, 207–213 (1923).
- Sturtevant, A. H. Inheritance of direction of coiling in *Limnaea*. *Science* **58**, 269–270 (1923).
- Freeman, G. & Lundelius, J. The developmental genetics of dextrality and sinistrality in the gastropod *Limnaea peregra*. *Wilhelm Roux Arch. Dev. Biol.* **191**, 69–83 (1982).
- Nederbragt, A. J., van Loon, A. E. & Dictus, W. J. Evolutionary biology: hedgehog crosses the snail's midline. *Nature* **417**, 811–812 (2002).
- van den Biggelaar, J. A. M., van Loon, A. E. & Damen, W. G. M. Mesentoblast and trophoblast specification in species with spiral cleavage predict their phyletic relations. *Neth. J. Zool.* **46**, 8–21 (1995).
- Lartillot, N., Lespinet, O., Vervoort, M. & Adoutte, A. Expression pattern of *Brachyury* in the mollusc *Patella vulgata* suggests a conserved role in the establishment of the AP axis in Bilateria. *Development* **129**, 1411–1421 (2002).

15. Dictus, W. J. A. G. & Damen, P. Cell lineage and clonal-contribution map of the trophophore larva of *Patella vulgata* (Mollusca). *Mech. Dev.* **62**, 213–226 (1997).
16. Inman, G. J. *et al.* SB-431542 is a potent and specific inhibitor of transforming growth factor- $\beta$  superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **62**, 65–74 (2002).
17. Christiaen, L. *et al.* Evolutionary modification of mouth position in deuterostomes. *Semin. Cell Dev. Biol.* **18**, 502–511 (2007).
18. Nogi, T., Yuan, Y. E., Sorocco, D., Perez-Tomas, R. & Levin, M. Eye regeneration assay reveals an invariant functional left-right asymmetry in the early bilaterian, *Dugesia japonica*. *L laterality* **10**, 193–205 (2005).
19. Oviedo, N. J. & Levin, M. Gap junctions provide new links in left-right patterning. *Cell* **129**, 645–647 (2007).
20. Ponder, W. F. & Lindberg, D. R. Towards a phylogeny of gastropod molluscs: analysis using morphological characters. *Zool. J. Linn. Soc.* **119**, 83–265 (1997).
21. Hibino, T., Nishino, A. & Amemiya, S. Phylogenetic correspondence of the body axes in bilaterians is revealed by the right-sided expression of Pitx genes in echinoderm larvae. *Dev. Growth Differ.* **48**, 587–595 (2006).
22. Palmer, A. R. Symmetry breaking and the evolution of development. *Science* **306**, 828–833 (2004).
23. Duboc, V. & Lepage, T. A conserved role for the Nodal signaling pathway in the establishment of dorso-ventral and left-right axes in deuterostomes. *J. Exp. Zool. B Mol. Dev. Evol.* **310**, 41–53 (2008).
24. Levin, M. Left-right asymmetry in embryonic development: a comprehensive review. *Mech. Dev.* **122**, 3–25 (2005).
25. Gould, M. C., Stephano, J. L., Ortiz-Barrón, B. J. & Pérez-Quezada, I. Maturation and fertilization in *Lottia gigantea* oocytes: intracellular pH, Ca<sup>2+</sup>, and electrophysiology. *J. Exp. Zool.* **290**, 411–420 (2001).
26. Price, A. L. & Patel, N. H. Investigating divergent mechanisms of mesoderm development in arthropods: the expression of Ph-twist and Ph-mef 2 in *Parhyale hawaiensis*. *J. Exp. Zoolog. B Mol. Dev. Evol.* **310**, 24–40 (2008).
27. Shibasaki, Y., Shimuzu, M. & Kuroda, R. Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. *Curr. Biol.* **14**, 1462–1467 (2004).
28. Camey, T. & Verdonk, N. H. The early development of the snail *Biomphalaria glabrata* (Say) and the origin of the head organs. *Neth. J. Zool.* **20**, 93–121 (1970).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank E. Begovic, E. E. Gonzales and I. Martínez-Solano for help collecting and fertilizing *L. gigantea*. A. Almeida helped with the drug experiments. D. R. Lindberg provided the *Busycon pulleyi* specimen. The NIAID Schistosomiasis Resource Center provided adults of *B. glabrata*. We thank M. Levine, D. R. Lindberg, P. Liu, M. Modrell, S. Nichols, M. Protas and J. Rehm for comments on the manuscript, and I. Hariharan for suggesting the experiments with rapamycin. C.G. was sponsored by a postdoctoral fellowship of the Ministerio de Educacion y Ciencia (Spain) and the Center for Integrative Genomics. N.H.P. is an Investigator of the Howard Hughes Medical Institute.

**Author Contributions** C.G. performed experiments; C.G. and N.H.P. designed experiments, collected and analysed data, and wrote the manuscript.

**Author Information** *Nodal*, *Pitx* and *hedgehog* sequences of *L. gigantea* and *B. glabrata* are deposited at the EMBL-GenBank data libraries; accession numbers EU394708 and EU394707 (for *nodal*), EU797117 and EU797116 (for *Pitx*) and EU394706 and EU394705 (for *hedgehog*). In addition, a *brachury* sequence of *L. gigantea* is deposited as accession number EU797118. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to N.H.P. (nipam@uclink.berkeley.edu).

# LETTERS

## Detecting influenza epidemics using search engine query data

Jeremy Ginsberg<sup>1</sup>, Matthew H. Mohebbi<sup>1</sup>, Rajan S. Patel<sup>1</sup>, Lynnette Brammer<sup>2</sup>, Mark S. Smolinski<sup>1</sup> & Larry Brilliant<sup>1</sup>

Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year<sup>1</sup>. In addition to seasonal influenza, a new strain of influenza virus against which no previous immunity exists and that demonstrates human-to-human transmission could result in a pandemic with millions of fatalities<sup>2</sup>. Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza<sup>3,4</sup>. One way to improve early detection is to monitor health-seeking behaviour in the form of queries to online search engines, which are submitted by millions of users around the world each day. Here we present a method of analysing large numbers of Google search queries to track influenza-like illness in a population. Because the relative frequency of certain queries is highly correlated with the percentage of physician visits in which a patient presents with influenza-like symptoms, we can accurately estimate the current level of weekly influenza activity in each region of the United States, with a reporting lag of about one day. This approach may make it possible to use search queries to detect influenza epidemics in areas with a large population of web search users.

Traditional surveillance systems, including those used by the US Centers for Disease Control and Prevention (CDC) and the European Influenza Surveillance Scheme (EISS), rely on both virological and clinical data, including influenza-like illness (ILI) physician visits. The CDC publishes national and regional data from these surveillance systems on a weekly basis, typically with a 1–2-week reporting lag.

In an attempt to provide faster detection, innovative surveillance systems have been created to monitor indirect signals of influenza activity, such as call volume to telephone triage advice lines<sup>5</sup> and over-the-counter drug sales<sup>6</sup>. About 90 million American adults are believed to search online for information about specific diseases or medical problems each year<sup>7</sup>, making web search queries a uniquely valuable source of information about health trends. Previous attempts at using online activity for influenza surveillance have counted search queries submitted to a Swedish medical website (A. Hult, G. Rydevik and A. Linde, manuscript in preparation), visitors to certain pages on a US health website<sup>8</sup>, and user clicks on a search keyword advertisement in Canada<sup>9</sup>. A set of Yahoo search queries containing the words ‘flu’ or ‘influenza’ were found to correlate with virological and mortality surveillance data over multiple years<sup>10</sup>.

Our proposed system builds on this earlier work by using an automated method of discovering influenza-related search queries. By processing hundreds of billions of individual searches from 5 years of Google web search logs, our system generates more comprehensive models for use in influenza surveillance, with regional and state-level estimates of ILI activity in the United States. Widespread global usage of online search engines may eventually enable models to be developed in international settings.

By aggregating historical logs of online web search queries submitted between 2003 and 2008, we computed a time series of weekly counts for 50 million of the most common search queries in the United States. Separate aggregate weekly counts were kept for every query in each state. No information about the identity of any user was retained. Each time series was normalized by dividing the count for each query in a particular week by the total number of online search queries submitted in that location during the week, resulting in a query fraction (Supplementary Fig. 1).

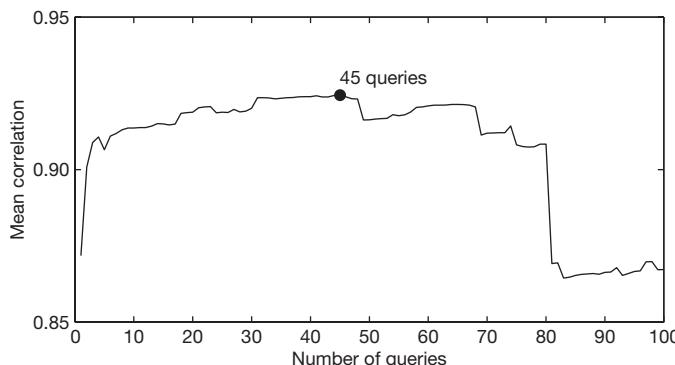
We sought to develop a simple model that estimates the probability that a random physician visit in a particular region is related to an ILI; this is equivalent to the percentage of ILI-related physician visits. A single explanatory variable was used: the probability that a random search query submitted from the same region is ILI-related, as determined by an automated method described below. We fit a linear model using the log-odds of an ILI physician visit and the log-odds of an ILI-related search query:  $\text{logit}(I(t)) = \alpha \text{logit}(Q(t)) + \varepsilon$ , where  $I(t)$  is the percentage of ILI physician visits,  $Q(t)$  is the ILI-related query fraction at time  $t$ ,  $\alpha$  is the multiplicative coefficient, and  $\varepsilon$  is the error term.  $\text{logit}(p)$  is simply  $\ln(p/(1-p))$ .

Publicly available historical data from the CDC’s US Influenza Sentinel Provider Surveillance Network (<http://www.cdc.gov/flu/weekly>) was used to help build our models. For each of the nine surveillance regions of the United States, the CDC reported the average percentage of all outpatient visits to sentinel providers that were ILI-related on a weekly basis. No data were provided for weeks outside of the annual influenza season, and we excluded such dates from model fitting, although our model was used to generate unvalidated ILI estimates for these weeks.

We designed an automated method of selecting ILI-related search queries, requiring no previous knowledge about influenza. We measured how effectively our model would fit the CDC ILI data in each region if we used only a single query as the explanatory variable,  $Q(t)$ . Each of the 50 million candidate queries in our database was separately tested in this manner, to identify the search queries which could most accurately model the CDC ILI visit percentage in each region. Our approach rewarded queries that showed regional variations similar to the regional variations in CDC ILI data: the chance that a random search query can fit the ILI percentage in all nine regions is considerably less than the chance that a random search query can fit a single location (Supplementary Fig. 2).

The automated query selection process produced a list of the highest scoring search queries, sorted by mean Z-transformed correlation across the nine regions. To decide which queries would be included in the ILI-related query fraction,  $Q(t)$ , we considered different sets of  $n$  top-scoring queries. We measured the performance of these models based on the sum of the queries in each set, and picked  $n$  such that we obtained the best fit against out-of-sample ILI data across the nine regions (Fig. 1).

<sup>1</sup>Google Inc., 1600 Amphitheatre Parkway, Mountain View, California 94043, USA. <sup>2</sup>Centers for Disease Control and Prevention, 1600 Clifton Road, NE, Atlanta, Georgia 30333, USA.



**Figure 1 | An evaluation of how many top-scoring queries to include in the ILI-related query fraction.** Maximal performance at estimating out-of-sample points during cross-validation was obtained by summing the top 45 search queries. A steep drop in model performance occurs after adding query 81, which is ‘oscar nominations’.

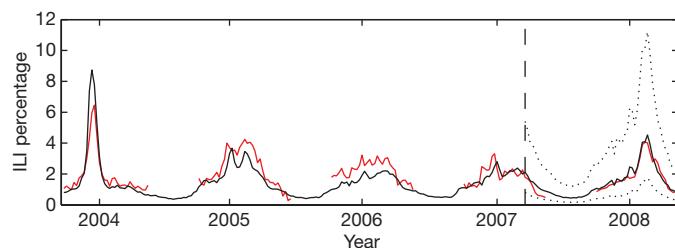
Combining the  $n = 45$  highest-scoring queries was found to obtain the best fit. These 45 search queries, although selected automatically, appeared to be consistently related to ILIs. Other search queries in the top 100, not included in our model, included topics like ‘high school basketball’, which tend to coincide with influenza season in the United States (Table 1).

Using this ILI-related query fraction as the explanatory variable, we fit a final linear model to weekly ILI percentages between 2003 and 2007 for all nine regions together, thus obtaining a single, region-independent coefficient. The model was able to obtain a good fit with CDC-reported ILI percentages, with a mean correlation of 0.90 (min = 0.80, max = 0.96,  $n = 9$  regions; Fig. 2).

The final model was validated on 42 points per region of previously untested data from 2007 to 2008, which were excluded from all previous steps. Estimates generated for these 42 points obtained a mean correlation of 0.97 (min = 0.92, max = 0.99,  $n = 9$  regions) with the CDC-observed ILI percentages.

Throughout the 2007–08 influenza season we used preliminary versions of our model to generate ILI estimates, and shared our results each week with the Epidemiology and Prevention Branch of Influenza Division at the CDC to evaluate timeliness and accuracy. Figure 3 illustrates data available at different points throughout the season. Across the nine regions, we were able to estimate consistently the current ILI percentage 1–2 weeks ahead of the publication of reports by the CDC’s US Influenza Sentinel Provider Surveillance Network.

Because localized influenza surveillance is particularly useful for public health planning, we sought to validate further our model



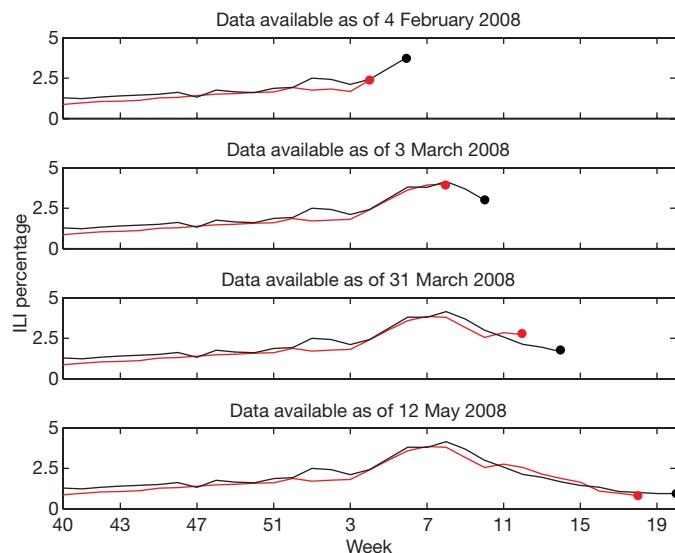
**Figure 2 | A comparison of model estimates for the mid-Atlantic region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated.** A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, whereas a correlation of 0.96 was obtained over 42 validation points. Dotted lines indicate 95% prediction intervals. The region comprises New York, New Jersey and Pennsylvania.

against weekly ILI percentages for individual states. The CDC does not make state-level data publicly available, but we validated our model against state-reported ILI percentages provided by the state of Utah, and obtained a correlation of 0.90 across 42 validation points (Supplementary Fig. 3).

Google web search queries can be used to estimate ILI percentages accurately in each of the nine public health regions of the United States. Because search queries can be processed quickly, the resulting ILI estimates were consistently 1–2 weeks ahead of CDC ILI surveillance reports. The early detection provided by this approach may become an important line of defence against future influenza epidemics in the United States, and perhaps eventually in international settings.

Up-to-date influenza estimates may enable public health officials and health professionals to respond better to seasonal epidemics. If a region experiences an early, sharp increase in ILI physician visits, it may be possible to focus additional resources on that region to identify the aetiology of the outbreak, providing extra vaccine capacity or raising local media awareness as necessary.

This system is not designed to be a replacement for traditional surveillance networks or supplant the need for laboratory-based diagnoses and surveillance. Notable increases in ILI-related search activity



**Figure 3 | ILI percentages estimated by our model (black) and provided by the CDC (red) in the mid-Atlantic region, showing data available at four points in the 2007–2008 influenza season.** During week 5 we detected a sharply increasing ILI percentage in the mid-Atlantic region; similarly, on 3 March our model indicated that the peak ILI percentage had been reached during week 8, with sharp declines in weeks 9 and 10. Both results were later confirmed by CDC ILI data.

**Table 1 | Topics found in search queries which were found to be most correlated with CDC ILI data**

Search query topic	Top 45 queries <i>n</i>	Top 45 queries Weighted	Next 55 queries <i>n</i>	Next 55 queries Weighted
Influenza complication	11	18.15	5	3.40
Cold/flu remedy	8	5.05	6	5.03
General influenza symptoms	5	2.60	1	0.07
Term for influenza	4	3.74	6	0.30
Specific influenza symptom	4	2.54	6	3.74
Symptoms of an influenza complication	4	2.21	2	0.92
Antibiotic medication	3	6.23	3	3.17
General influenza remedies	2	0.18	1	0.32
Symptoms of a related disease	2	1.66	2	0.77
Antiviral medication	1	0.39	1	0.74
Related disease	1	6.66	3	3.77
Unrelated to influenza	0	0.00	19	28.37
Total	45	49.40	55	50.60

The top 45 queries were used in our final model; the next 55 queries are presented for comparison purposes. The number of queries in each topic is indicated, as well as query-volume-weighted counts, reflecting the relative frequency of queries in each topic.

may indicate a need for public health inquiry to identify the pathogen or pathogens involved. Demographic data, often provided by traditional surveillance, cannot be obtained using search queries.

In the event that a pandemic-causing strain of influenza emerges, accurate and early detection of ILI percentages may enable public health officials to mount a more effective early response. Although we cannot be certain how search engine users will behave in such a scenario, affected individuals may submit the same ILI-related search queries used in our model. Alternatively, panic and concern among healthy individuals may cause a surge in the ILI-related query fraction and exaggerated estimates of the ongoing ILI percentage.

The search queries in our model are not, of course, exclusively submitted by users who are experiencing influenza-like symptoms, and the correlations we observe are only meaningful across large populations. Despite strong historical correlations, our system remains susceptible to false alerts caused by a sudden increase in ILI-related queries. An unusual event, such as a drug recall for a popular cold or flu remedy, could cause such a false alert.

Harnessing the collective intelligence of millions of users, Google web search logs can provide one of the most timely, broad-reaching influenza monitoring systems available today. Whereas traditional systems require 1–2 weeks to gather and process surveillance data, our estimates are current each day. As with other syndromic surveillance systems, the data are most useful as a means to spur further investigation and collection of direct measures of disease activity. This system will be used to track the spread of ILI throughout the 2008–09 influenza season in the United States. Results are freely available online at <http://www.google.org/flutrends>.

## METHODS SUMMARY

**Privacy.** None of the queries in the Google database for this project can be associated with a particular individual. The database retains no information about the identity, internet protocol (IP) address, or specific physical location of any user. Furthermore, any original web search logs older than 9 months are being made anonymous in accordance with Google's privacy policy (<http://www.google.com/privacypolicy.html>).

**Search query database.** For the purposes of our database, a search query is a complete, exact sequence of terms issued by a Google search user; we don't combine linguistic variations, synonyms, cross-language translations, misspellings, or subsequences, although we hope to explore these options in future work. For example, we tallied the search query 'indications of flu' separately from the search queries 'flu indications' and 'indications of the flu'.

Our database of queries contains 50 million of the most common search queries on all possible topics, without pre-filtering. Billions of queries occurred infrequently and were excluded. Using the internet protocol address associated with each search query, the general physical location from which the query

originated can often be identified, including the nearest major city if within the United States.

**Model data.** In the query selection process, we fit per-query models using all weeks between 28 September 2003 and 11 March 2007 (inclusive) for which the CDC reported a non-zero ILI percentage, yielding 128 training points for each region (each week is one data point). Forty-two additional weeks of data (18 March 2007 through to 11 May 2008) were reserved for final validation. Search query data before 2003 was not available for this project.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 14 August; accepted 13 November 2008.**

**Published online 19 November 2008; corrected 19 February 2009 (details online).**

1. World Health Organization. Influenza fact sheet. <<http://www.who.int/mediacentre/factsheets/2003/fs211/en/>> (2003).
2. World Health Organization. WHO consultation on priority public health interventions before and during an influenza pandemic. <[http://www.who.int/csr/disease/avian\\_influenza/consultation/en/](http://www.who.int/csr/disease/avian_influenza/consultation/en/)> (2004).
3. Ferguson, N. M. et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214 (2005).
4. Longini, I. M. et al. Containing pandemic influenza at the source. *Science* **309**, 1083–1087 (2005).
5. Espino, J., Hogan, W. & Wagner, M. Telephone triage: A timely data source for surveillance of influenza-like diseases. *AMIA Annu. Symp. Proc.* 215–219 (2003).
6. Magruder, S. Evaluation of over-the-counter pharmaceutical sales as a possible early warning indicator of human disease. *Johns Hopkins APL Tech. Digest* **24**, 349–353 (2003).
7. Fox, S. Online Health Search 2006. Pew Internet & American Life Project <[http://www.pewinternet.org/pdfs/pip\\_online\\_health\\_2006.pdf](http://www.pewinternet.org/pdfs/pip_online_health_2006.pdf)> (2006).
8. Johnson, H. et al. Analysis of Web access logs for surveillance of influenza. *Stud. Health Technol. Inform.* **107**, 1202–1206 (2004).
9. Eysenbach, G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu. Symp. Proc.* 244–248 (2006).
10. Polgreen, P. M., Chen, Y., Pennock, D. M. & Forrest, N. D. Using internet searches for influenza surveillance. *Clin. Infect. Dis.* **47**, 1443–1448 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank L. Finelli for providing background knowledge, helping us validate results and comments on this manuscript. We are grateful to R. Rolfs, L. Wyman and M. Patton for providing ILI data. We thank V. Sahai for his contributions to data collection and processing, and C. Nevill-Manning, A. Roetter and K. Sarvian for their comments on this manuscript.

**Author Contributions** J.G. and M.H.M. conceived, designed and implemented the system. J.G., M.H.M. and R.S.P. analysed the results and wrote the paper. L.B. contributed data. All authors edited and commented on the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to J.G. or M.H.M. (flutrends-support@google.com).

## METHODS

**Automated query selection process.** Using linear regression with fourfold cross validation, we fit models to four 96-point subsets of the 128 points in each region. Each per-query model was validated by measuring the correlation between the model's estimates for the 32 held-out points and the CDC's reported regional ILI percentage at those points. Temporal lags were considered, but ultimately not used in our modelling process.

Each candidate search query was evaluated nine times, once per region, using the search data originating from a particular region to explain the ILI percentage in that region. With four cross-validation folds per region, we obtained 36 different correlations between the candidate model's estimates and the observed ILI percentages. To combine these into a single measure of the candidate query's performance, we applied the Fisher Z-transformation<sup>11</sup> to each correlation, and took the mean of the 36 Z-transformed correlations.

**Computation and pre-filtering.** In total, we fit 450 million different models to test each of the candidate queries. We used a distributed computing framework<sup>12</sup> to divide the work among hundreds of machines efficiently. The amount of computation required could have been reduced by making assumptions about which queries might be correlated with ILI. For example, we could have attempted to eliminate non-influenza-related queries before fitting any models. However, we were concerned that aggressive filtering might accidentally eliminate valuable data. Furthermore, if the highest-scoring queries seemed entirely unrelated to influenza, it would provide evidence that our query selection approach was invalid.

**Constructing the ILI-related query fraction.** We concluded the query selection process by choosing to keep the search queries whose models obtained the highest mean Z-transformed correlations across regions: these queries were deemed to be 'ILI-related'.

To combine the selected search queries into a single aggregate variable, we summed the query fractions on a regional basis, yielding our estimate of the ILI-related query fraction,  $Q(t)$ , in each region. Note that the same set of queries was selected for each region.

**Fitting and validating a final model.** We fit one final univariate model, used for making estimates in any region or state based on the ILI-related query fraction from that region or state. We regressed over 1,152 points, combining all 128 training points used in the query selection process from each of the nine regions. We validated the accuracy of this final model by measuring its performance on 42 additional weeks of previously untested data in each region, from the most recently available time period (18 March 2007 through to 11 May 2008). These 42 points represent approximately 25% of the total data available for the project, the first 75% of which was used for query selection and model fitting.

**State-level model validation.** To evaluate the accuracy of state-level ILI estimates generated using our final model, we compared our estimates against weekly ILI percentages provided by the state of Utah. Because the model was fit using regional data through 11 March 2007, we validated our Utah ILI estimates using 42 weeks of previously untested data, from the most recently available time period (18 March 2007 through to 11 May 2008).

11. David, F. The moments of the z and F distributions. *Biometrika* **36**, 394–403 (1949).
12. Dean, J. & Ghemawat, S. Mapreduce: Simplified data processing on large clusters. *Sixth Symp. Oper. Syst. Des. Implement.*, (2004).

## LETTERS

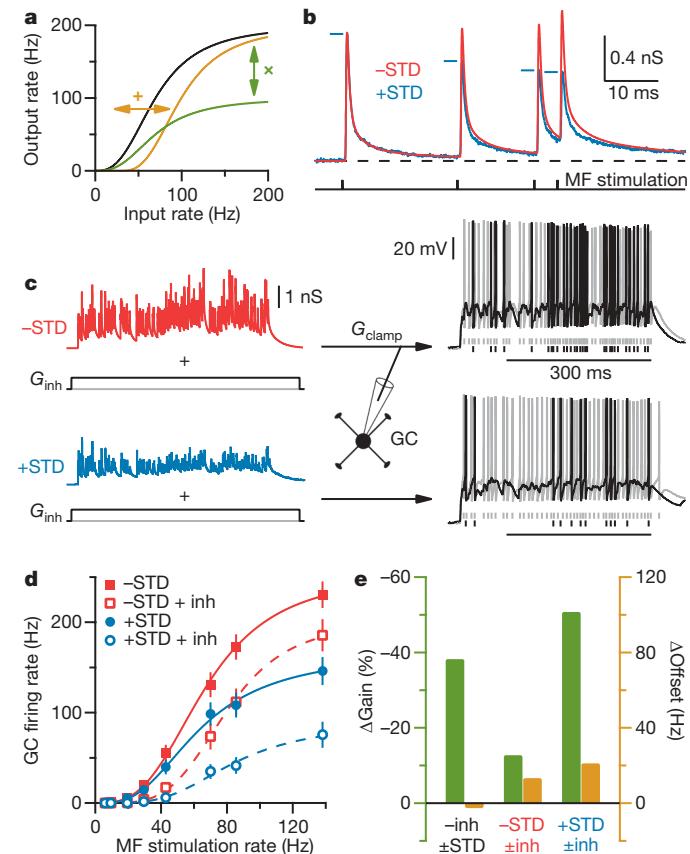
# Synaptic depression enables neuronal gain control

Jason S. Rothman<sup>1</sup>, Laurence Cathala<sup>1\*</sup>, Volker Steuber<sup>1\*</sup> & R. Angus Silver<sup>1</sup>

To act as computational devices, neurons must perform mathematical operations as they transform synaptic and modulatory input into output firing rate<sup>1</sup>. Experiments and theory indicate that neuronal firing typically represents the sum of synaptic inputs<sup>1–3</sup>, an additive operation, but multiplication of inputs is essential for many computations<sup>1</sup>. Multiplication by a constant produces a change in the slope, or gain, of the input–output relationship, amplifying or scaling down the sensitivity of the neuron to changes in its input. Such gain modulation occurs *in vivo*, during contrast invariance of orientation tuning<sup>4</sup>, attentional scaling<sup>5</sup>, translation-invariant object recognition<sup>6</sup>, auditory processing<sup>7</sup> and coordinate transformations<sup>8,9</sup>. Moreover, theoretical studies highlight the necessity of gain modulation in several of these tasks<sup>9–11</sup>. Although potential cellular mechanisms for gain modulation have been identified, they often rely on membrane noise and require restrictive conditions to work<sup>3,12–18</sup>. Because nonlinear components are used to scale signals in electronics, we examined whether synaptic nonlinearities are involved in neuronal gain modulation. We used synaptic stimulation and the dynamic-clamp technique to investigate gain modulation in granule cells in acute slices of rat cerebellum. Here we show that when excitation is mediated by synapses with short-term depression (STD), neuronal gain is controlled by an inhibitory conductance in a noise-independent manner, allowing driving and modulatory inputs to be multiplied together. The nonlinearity introduced by STD transforms inhibition-mediated additive shifts in the input–output relationship into multiplicative gain changes. When granule cells were driven with bursts of high-frequency mossy fibre input, as observed *in vivo*<sup>19,20</sup>, larger inhibition-mediated gain changes were observed, as expected with greater STD. Simulations of synaptic integration in more complex neocortical neurons suggest that STD-based gain modulation can also operate in neurons with large dendritic trees. Our results establish that neurons receiving depressing excitatory inputs can act as powerful multiplicative devices even when integration of postsynaptic conductances is linear.

The way a neuron transforms signals can be captured by its input–output relationship (Fig. 1a). A modulatory input can change the shape of this relationship, thereby performing a mathematical operation on this transfer function. A shift along the input axis corresponds to an additive operation (+), whereas a change in slope corresponds to a multiplicative operation, or gain change (×). Cerebellar granule cells (GCs) are well suited for studying gain modulation because they have few synaptic inputs. Excitation comes from approximately four mossy fibres (MFs), which can sustain rate-coded signals over a large bandwidth<sup>20</sup> and have STD<sup>19,21</sup>. Inhibition comes from Golgi cells, most of which is mediated by a modulatable tonic GABA<sub>A</sub> ( $\gamma$ -aminobutyric acid subtype A) receptor (GABAR) conductance<sup>22</sup>. Because it is difficult to activate multiple inputs independently, and because we wanted to compare real synaptic inputs showing frequency-dependent STD with artificial synaptic inputs without STD, we used dynamic-clamp to study synaptic integration.

GCs are ideal for this because their soma and dendrites form a single electrical compartment, allowing dendritic inputs to be mimicked by somatic conductance injection<sup>13</sup>.

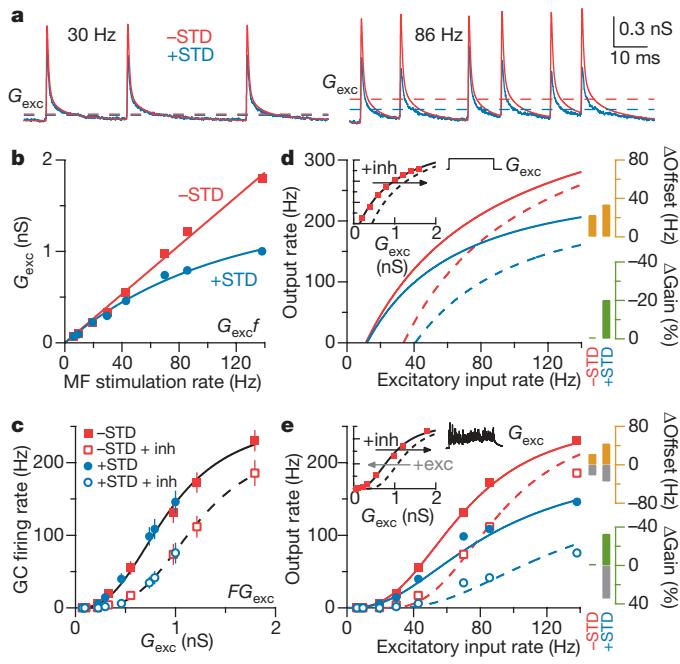


**Figure 1 | Synaptic depression enhances inhibition-mediated gain modulation.** **a**, Hypothetical neuronal input–output relationship before (black) and after multiplicative gain modulation (green,  $\times$ ) and an additive offset (orange, +). **b**, Averaged AMPAR-mediated synaptic train showing STD (blue trace) in response to Poisson stimulation (black ticks;  $f = 86$  Hz) of single MF inputs; blue ticks show the peak amplitude of the depressing synaptic conductances. The red trace shows the corresponding artificial synaptic train without STD. The black dashed line is the zero conductance level. **c**, The sum of four independent synaptic trains (each  $f = 86$  Hz) with and without STD injected into a GC by means of dynamic clamp ( $G_{\text{clamp}}$ ) with and without tonic inhibition (black and grey, respectively;  $G_{\text{inh}} = 500$  pS). Right vertical ticks indicate spike times. Horizontal bars indicate output rate measurement window.  $V_{\text{rest}} = -79$  mV. **d**, Average input–output relationships ( $n = 9$ ) with and without STD (blue and red, respectively) and tonic inhibition (inh, open and filled symbols, respectively; error bars are  $\pm$ s.e.m.). Lines are fits to a Hill function (equation (5) in Methods; Supplementary Table 1). **e**, Gain (green) and offset (orange) changes due to STD ( $\pm$ STD) and inhibition ( $\pm$ inh) from fits in **d**.

<sup>1</sup>Department of Neuroscience, Physiology and Pharmacology, University College London, Gower Street, London WC1E 6BT, UK.

\*These authors contributed equally to this work.

We measured AMPA ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor (AMPAR)-mediated excitatory post-synaptic currents (EPSCs) from mature GCs at physiological temperature during Poisson-type stimulation of single MFs at different frequencies. These EPSC trains, which showed STD, were then converted to conductance (Fig. 1b, +STD, blue traces throughout). Artificial conductance trains without STD (−STD, red traces throughout) but with identical event timing were constructed by adding a fixed amplitude synaptic conductance waveform at each stimulus time. The effects of STD on synaptic integration were then investigated by injecting the sum of four statistically independent conductance trains with the same mean MF rate into GCs using dynamic-clamp and measuring the mean output firing rate with and without STD (Fig. 1c). The resulting steady-state input–output relationships revealed that STD had an almost purely multiplicative effect under control conditions (Fig. 1d). Multiplicative and additive transformations of the input–output relationship were quantified by fitting the data to Hill-like equations<sup>16</sup> (Fig. 1d, equation (5) in Methods and Supplementary Table 1) and measuring the change in slope (Fig. 1e,  $\Delta$ Gain, green) and the shift in the half maximal response ( $\Delta$ Offset, orange). In the absence of STD, adding a 500 pS tonic inhibitory conductance ( $G_{inh}$ ; Fig. 1c) close to the physiological value ( $438 \pm 93$  pS (mean  $\pm$  s.e.m.);  $n = 10$ ) produced a modest scaling of the input–output relationship together with an additive shift<sup>13</sup> (Fig. 1d, e). In contrast, the same level of inhibition produced a fourfold larger gain change in the input–output relationship when driven with depressing synapses. Moreover, this multiplicative scaling was nearly constant over the entire input range (Supplementary Fig. 1).

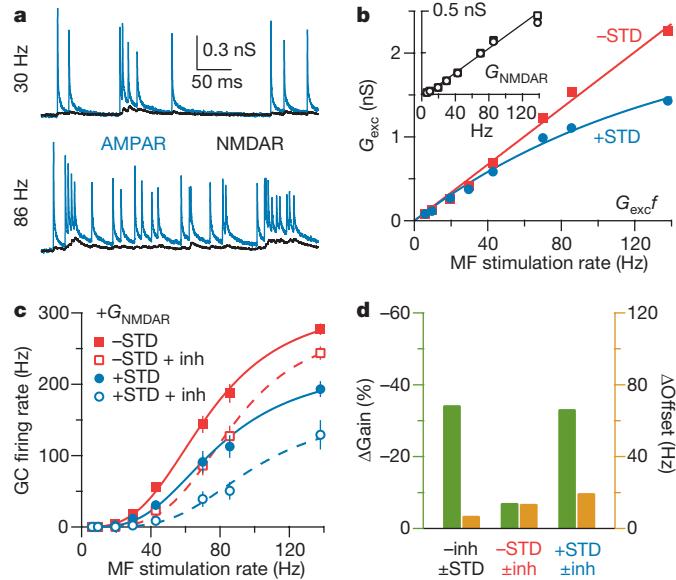


**Figure 2 | Synaptic depression transforms additive shifts into multiplicative gain modulation.** **a**, Average AMPAR-mediated synaptic trains with and without STD (blue and red, respectively) with their time-averaged conductances ( $G_{exc}$ , dashed lines). **b**,  $G_{exc}$  versus MF rate ( $f$ ) with and without STD. The red line is a linear fit (equation (3), Methods); the blue line is an exponential fit (equation (4)). **c**, Data in Fig. 1d are plotted as output rate ( $F$ ) versus  $G_{exc}$ . Lines are Hill fits (equation (5)) and error bars are  $\pm$  s.e.m. **d**, Predicted input–output relationships computed from  $G_{exc}f$  relationships in **b**, and experimentally measured GC  $FG_{exc}$  relationship generated with noise-free excitatory conductance steps (inset, red squares from Fig. 2b in ref. 13). The solid line in inset is a Hill fit, and the dashed line is the same curve rightward shifted 0.3 nS along the  $G_{exc}$  axis. Bar graphs show changes in gain (green) and offset (orange) due to inhibition with and without STD. **e**, Same as **d**, except control  $FG_{exc}$  data (inset) are from **c** during noisy synaptic excitation. The predicted change in the input–output relationship for STD match the experimental observations (blue symbols). Grey bars show gain and offset changes for a leftward shift of 0.3 nS to simulate excitation.

To understand how STD performs multiplicative gain modulation, we examined the frequency-dependence of the time-averaged excitatory conductances ( $G_{exc}$ ; Fig. 2a, dashed lines). Without STD, the relationship between  $G_{exc}$  and MF input rate ( $f$ ) was linear (Fig. 2b). With STD, it became noticeably sublinear above 40 Hz and could be fit with a saturating exponential function (Fig. 2b). To examine how  $G_{exc}$  was integrated postsynaptically, we plotted the mean GC firing rate ( $F$ ) as a function of  $G_{exc}$ . We found that the  $FG_{exc}$  relationships with and without STD overlaid (Fig. 2c), indicating that  $G_{exc}$  was integrated similarly for both cases. Thus, the nonlinearity between  $f$  and  $G_{exc}$  introduced by STD underlies the enhanced gain modulation.

Because neuronal gain can be altered by synaptic noise<sup>3,13</sup>, we examined how STD transforms GC  $FG_{exc}$  relationships generated with noise-free conductance steps (Fig. 2d, inset<sup>13</sup>) and noisy synaptic conductances (Fig. 2e, inset). Hill equations were used to mimic  $FG_{exc}$  relationships (Fig. 2d, e, insets) whereas linear and exponential functions were used to represent  $G_{exc}f$  relationships (Fig. 2b). These were then combined to predict the input–output relationships. In the absence of STD, additive shifts in the  $FG_{exc}$  relationship produced a purely additive shift in the input–output relationship (Fig. 2d, e), consistent with conductance performing an additive operation when the noise level is constant<sup>2,3</sup>. In contrast, when STD was present, additive shifts in the  $FG_{exc}$  relationship produced robust gain changes in the input–output relationship for both noise-free and noisy excitation (Fig. 2d, e). Indeed, the gain change predicted from purely additive shifts in the  $FG_{exc}$  relationship accounted for most of that observed experimentally (Fig. 2e). This shows that STD transforms linear, additive modulation of the  $FG_{exc}$  relationship into multiplicative gain modulation of the input–output relationship, irrespective of the presence of synaptic noise.

Our experimental results were reproduced in a conductance-based integrate-and-fire model that included short-term synaptic plasticity (Supplementary Fig. 2a, b) confirming the primary role of STD in gain modulation, rather than synaptic noise or other nonlinear synaptic

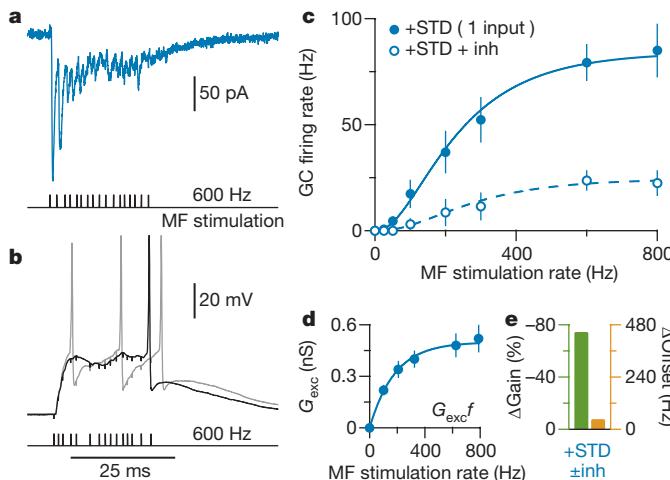


**Figure 3 | Gain modulation in the presence of synaptic NMDAR conductances.** **a**, Average AMPAR (blue) and NMDAR (black) synaptic conductance trains from the same four GCs during Poisson stimulation of individual MF inputs. **b**,  $G_{exc}$  versus MF rate ( $f$ ) for AMPAR + NMDAR components with and without AMPAR STD. Red and blue lines are fits to equations (3) and (4) (Methods), respectively. Inset:  $G_{exc}f$  relationships for NMDAR with and without AMPAR STD (circles and squares, respectively). **c**, Average input–output relationships for AMPAR + NMDAR with and without AMPAR STD and tonic inhibition. Lines are Hill fits (equation (5)) and error bars are  $\pm$  s.e.m. **d**, Gain (green) and offset (orange) changes due to STD ( $\pm$ STD) and tonic inhibition ( $\pm$ inh) from fits in **c**.

mechanisms, such as AMPAR activation by means of spillover<sup>23</sup>. Increasing the level of STD increased the inhibition-mediated gain reduction (Supplementary Fig. 2c, d). Conversely, increasing the level of tonic inhibition increased the size of the gain reduction at all levels of STD. We also observed gain modulation with phasic inhibitory conductances, which are widespread in the central nervous system (Supplementary Fig. 3), confirming that STD-mediated gain modulation operates effectively with both phasic and tonic inhibition.

NMDA (*N*-methyl-D-aspartate) receptors (NMDARs) contribute to transmission at many central synapses and show a nonlinear voltage dependence. To examine experimentally how this synaptic component interacts with STD-based gain modulation, we measured the NMDAR component during Poisson stimulation of a single MF input and added it to the AMPAR conductance trains from the same GCs (Fig. 3a). The  $G_{\text{exc}}$  $f$  relationships for the combined AMPAR and NMDAR components (Fig. 3b) were similar in shape to the AMPAR-only cases with and without STD (Fig. 2b), because the relationship between the mean NMDAR conductance and MF frequency was approximately linear (Fig. 3b, inset, and Supplementary Fig. 4). In the absence of AMPAR STD, 500 pS inhibition produced a slightly smaller gain reduction in the GC input–output relationship (Fig. 3c, d, 7.1%) than the AMPAR component alone (Fig. 1d, e, 12.6%), as predicted for noise-based gain modulation<sup>24</sup>. However, a robust inhibition-mediated gain reduction was observed with the depressing AMPAR component (Fig. 3d, 33.1%). These results show that synaptic activation of NMDARs in adult GCs, which contain the NR2C subunit<sup>25</sup>, is approximately linear with frequency and that this has little effect on STD-based gain modulation.

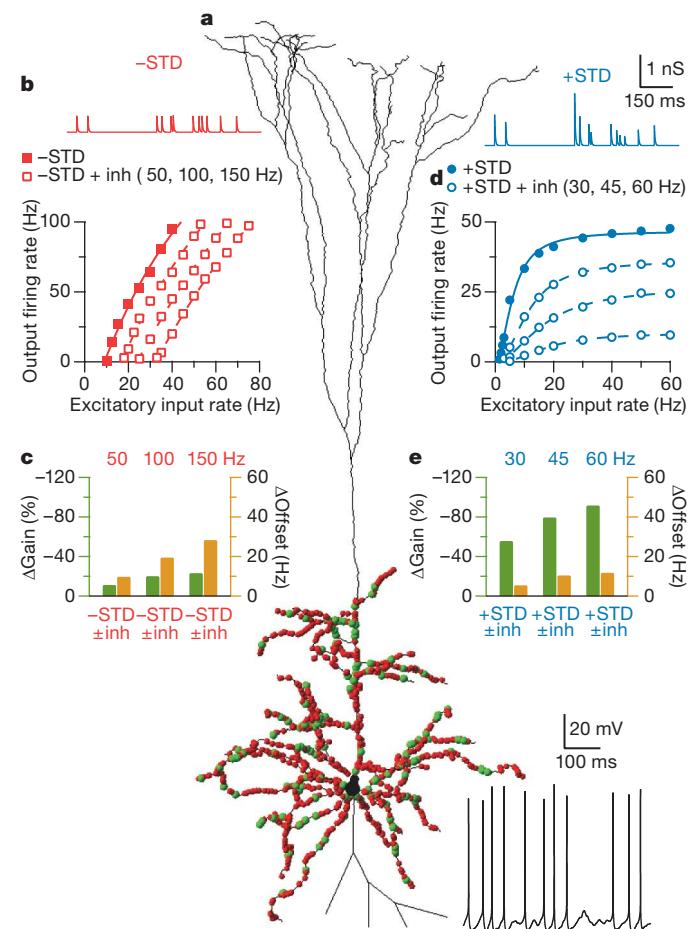
*In vivo* recordings show that finger extension<sup>20</sup> and facial stimulation<sup>19</sup> can produce high-frequency bursts of MF firing. To examine whether STD-mediated gain modulation can operate under such conditions, we recorded from GCs while stimulating individual MFs with high-frequency bursts to mimic activity *in vivo*<sup>19,20</sup>. This induced bursts of mixed AMPAR–NMDAR EPSCs under voltage-clamp, confirming reliable MF activation (Fig. 4a). In current-clamp, this produced GC firing from a potential of  $-75$  mV (Fig. 4b), as observed *in vivo*<sup>19</sup>. Firing rate was measured from the fourth stimulus to allow



**Figure 4 | Gain modulation during broad bandwidth single mossy fibre stimulation.** **a**, Mixed AMPAR + NMDAR EPSCs in GCs during single MF, Poisson-burst-like stimulation (black ticks) at  $-75$  mV. Stimulus artefacts were subtracted. **b**, Voltage responses to burst-like MF stimulation with and without tonic inhibition (black and grey, respectively;  $G_{\text{inh}} = 500$  pS) injected by means of dynamic clamp. The black horizontal bar indicates the firing rate measurement window.  $V_{\text{rest}} = -74$  mV. **c**, Average GC input–output relationship with and without inhibition for single MF stimulation ( $n = 7$ ) with error bars ( $\pm$  s.e.m.) and Hill fits (equation (5)). **d**, Mean relationship between time-averaged conductance ( $G_{\text{exc}}$ ) and MF stimulation rate ( $f$ ). The blue line is an exponential fit (equation (4)). **e**, Gain (green) and offset (orange) changes due to inhibition ( $\pm$  inh) computed from fits in **c**.

time for EPSC depression to occur<sup>21</sup>. A 500 pS tonic inhibitory conductance reduced firing (Fig. 4b) and produced a robust reduction in neuronal gain (Fig. 4c) that was twofold larger than for the four-fibre excitation (Figs 3d and 4e). The powerful nonlinearity between  $G_{\text{exc}}$  and  $f$  (Fig. 4d), combined with a purely additive shift in the  $G_{\text{exc}}$  $f$  relationship, accounted for most of the inhibition-mediated gain change (Supplementary Fig. 5), confirming that it was mediated predominantly by STD.

Several factors complicate the implementation of multiplicative operations in neurons with extensive dendritic trees. The large voltage fluctuations required for noise-based gain modulation<sup>3,13,16,17</sup> are difficult to achieve with many inputs and heavy dendritic filtering, without precisely balanced excitation and inhibition<sup>3</sup>. Moreover, multiplication by means of nonlinear dendritic integration<sup>26</sup> or a combination of noise and dendritic saturation<sup>17</sup> requires spatially localized activation of synapses on a single dendritic branch, but these are usually distributed over the tree, making linear synaptic integration more likely<sup>26</sup>. To test whether the STD-mediated gain modulation can operate effectively in more complex neurons than GCs, we simulated synaptic integration in a neocortical layer 5 neuron model<sup>27</sup> with hundreds of synaptic inputs, randomly distributed over the



**Figure 5 | Multiplicative gain modulation in a cortical layer 5 pyramidal neuron model.** **a**, Layer 5 neuron model with location of excitatory (red circles) and inhibitory (green circles) synaptic contacts. Bottom: black trace shows spiking during synaptic excitation and inhibition. **b**, Top: conductance train for non-depressing unitary excitatory synaptic input. Bottom: input–output relationship for non-depressing synaptic excitation ( $-STD$ , red) for control conditions (solid symbols) and various rates of synaptic inhibition (open symbols) together with Hill fits (equation (5), Methods). **c**, Gain (green) and offset (orange) changes due to different inhibition rates ( $\pm$  inh) in the absence of STD from fits in **b**. **d**, As for **b**, but for depressing synaptic excitation ( $+STD$ , blue) together with Hill fits. **e**, As for **c**, but for depressing synaptic excitation ( $+STD$ ) from fits in **d**.

basolateral dendritic tree (Fig. 5a and Methods). These simulations show that inhibitory synaptic input can act as a powerful modulator of neuronal gain when excitation is mediated by depressing synapses and yet perform a largely additive operation when it is absent (Fig. 5b–e). These results indicate that multiplicative operations are possible in cells with large dendritic arbors under conditions of linear integration when excitation is mediated by depressing synapses.

*In vivo*, MF–GC synapses typically operate at high frequencies<sup>20</sup>, where STD is pronounced<sup>19</sup>, although low-frequency vestibular inputs are a notable exception, where STD is absent and linear transmission is preserved<sup>28</sup>. At 100 Hz, STD is mediated predominantly by AMPAR desensitization<sup>21</sup>. AMPARs can therefore act as nonlinear ‘molecular amplifiers’ that contribute to single-cell computation. Although vesicle replenishment is rapid at MF release sites<sup>21</sup>, it will become limiting at higher frequencies and presynaptic STD will dominate. Presynaptic STD has a range of frequency dependencies across central synapses and can be modulated by long-term plasticity<sup>29</sup>, potentially allowing gain modulation to be matched to the operational frequency of the inputs.

Multiplication of driving and modulatory input conductance is most effective when synaptic depression has reached steady state. Because EPSC depression occurs rapidly in GCs ( $\tau_{\text{onset}} = 1.5$  pulses)<sup>21</sup>, gain modulation will only require a 4–75-ms settling time ( $2\tau_{\text{onset}}/\text{frequency}$ ; 40–800 Hz), but at cortical synapses where STD occurs at lower frequencies gain modulation will be slower. STD-mediated gain modulation allows a modulatory conductance to scale a neuron’s sensitivity to all of its driving inputs. This is distinct from the transient synaptic responses to changes in input rate, which allows signalling independent of absolute rate<sup>30</sup>. STD can therefore perform distinct multiplicative operations on different time scales: transient enhancement of the gain of dynamically changing inputs before STD has developed, and multiplication of all driving and modulatory inputs once depression has occurred.

Our findings show that the mathematical operation performed by a modulatory input on a particular set of driving inputs depends on their STD characteristics. This could allow input-configuration-specific, and thus context-dependent, computation at the level of a single neuron. The widespread incidence of STD, its compatibility with linear synaptic integration and the lack of dependence on synaptic noise indicate that this cellular mechanism for gain modulation could be used to multiply inputs together in many different neuronal types in the brain.

## METHODS SUMMARY

Whole-cell recordings were made from GCs at  $35.3 \pm 0.1^\circ\text{C}$  ( $n = 62$ ) in rat cerebellar slices (postnatal day 30–40). EPSC trains for conductance clamp were recorded during stimulation of a single MF input at random Poisson intervals (mean rate  $f = 6$ –138 Hz). The NMDA EPSC component was isolated by recording in  $5\text{ }\mu\text{M}$  2,3-dioxo-6-nitro-1,2,3,4-tetrahydrobenzo[f]quinoxaline-7-sulphonamide (NBQX). The AMPAR EPSC component was obtained by subtracting control and NBQX recordings. Conductance trains at each frequency were constructed by averaging ten responses, subtracting stimulus artefacts, dividing by the driving force, and averaging responses from four different GCs. AMPAR, NMDA and tonic GABAR conductances with reversal potentials of 0 mV, 0 mV and  $-75$  mV, respectively, were injected into GCs using a dynamic-clamp amplifier. Inhibition was blocked with  $10\text{ }\mu\text{M}$  SR95531. Direct activation of GC firing was achieved with single MF excitation using a 15-pulse Poisson stimulation (25–800 Hz). To minimize the effects of any time-dependent changes in excitability, MF excitation was randomized, control and tonic inhibition measurements interleaved, and experiments were carried out within 15 min of obtaining intracellular access to the GC. The cortical layer 5 neuron model<sup>27</sup> was connected to 400 excitatory neurons and 30 inhibitory neurons using neuroConstruct (<http://www.NeuroConstruct.org>) and run on the NEURON simulator (<http://www.neuron.yale.edu>). Data are presented as mean  $\pm$  s.e.m.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 18 July; accepted 30 October 2008.

Published online 14 January 2009.

- Koch, C. *Biophysics of Computation: Information Processing in Single Neurons* 552 (Oxford Univ. Press, 1999).

- Holt, G. R. & Koch, C. Shunting inhibition does not have a divisive effect on firing rates. *Neural Comput.* 9, 1001–1013 (1997).
- Chance, F. S., Abbott, L. F. & Reyes, A. D. Gain modulation from background synaptic input. *Neuron* 35, 773–782 (2002).
- Anderson, J. S., Lampl, I., Gillespie, D. C. & Ferster, D. The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* 290, 1968–1972 (2000).
- Treue, S. & Martinez Trujillo, J. C. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399, 575–579 (1999).
- Tovee, M. J., Rolls, E. T. & Azzopardi, P. Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *J. Neurophysiol.* 72, 1049–1060 (1994).
- Ingham, N. J. & McAlpine, D. GABAergic inhibition controls neural gain in inferior colliculus neurons sensitive to interaural time differences. *J. Neurosci.* 25, 6187–6198 (2005).
- Brotchie, P. R., Andersen, R. A., Snyder, L. H. & Goodman, S. J. Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature* 375, 232–235 (1995).
- Yakusheva, T. A. et al. Purkinje cells in posterior cerebellar vermis encode motion in an inertial reference frame. *Neuron* 54, 973–985 (2007).
- Salinas, E. & Abbott, L. F. Transfer of coded information from sensory to motor networks. *J. Neurosci.* 15, 6461–6474 (1995).
- Pouget, A. & Sejnowski, T. J. Spatial transformations in the parietal cortex using basis functions. *J. Cogn. Neurosci.* 9, 222–237 (1997).
- Hansel, D. & van Vreeswijk, C. How noise contributes to contrast invariance of orientation tuning in cat visual cortex. *J. Neurosci.* 22, 5118–5128 (2002).
- Mitchell, S. J. & Silver, R. A. Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron* 38, 433–445 (2003).
- Gabbiani, F., Krapp, H. G., Koch, C. & Laurent, G. Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420, 320–324 (2002).
- Tiesinga, P. H., Jose, J. V. & Sejnowski, T. J. Comparison of current-driven and conductance-driven neocortical model neurons with Hodgkin-Huxley voltage-gated channels. *Phys. Rev. E* 62, 8413–8419 (2000).
- Murphy, B. K. & Miller, K. D. Multiplicative gain changes are induced by excitation or inhibition alone. *J. Neurosci.* 23, 10040–10051 (2003).
- Prescott, S. A. & De Koninck, Y. Gain control of firing rate by shunting inhibition: roles of synaptic noise and dendritic saturation. *Proc. Natl. Acad. Sci. USA* 100, 2076–2081 (2003).
- Fellous, J. M., Rudolph, M., Destexhe, A. & Sejnowski, T. J. Synaptic background noise controls the input/output characteristics of single cells in an *in vitro* model of *in vivo* activity. *Neuroscience* 122, 811–829 (2003).
- Rancz, E. A. et al. High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. *Nature* 450, 1245–1248 (2007).
- van Kan, P. L., Gibson, A. R. & Houk, J. C. Movement-related inputs to intermediate cerebellum of the monkey. *J. Neurophysiol.* 69, 74–94 (1993).
- Saviane, C. & Silver, R. A. Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse. *Nature* 439, 983–987 (2006).
- Semyanov, A., Walker, M. C., Kullmann, D. M. & Silver, R. A. Tonically active GABA A receptors: modulating gain and maintaining the tone. *Trends Neurosci.* 27, 262–269 (2004).
- DiGregorio, D. A., Nusser, Z. & Silver, R. A. Spillover of glutamate onto synaptic AMPA receptors enhances fast transmission at a cerebellar synapse. *Neuron* 35, 521–533 (2002).
- Berends, M., Maex, R. & De Schutter, E. The effect of NMDA receptors on gain modulation. *Neural Comput.* 17, 2531–2547 (2005).
- Cathala, L., Misra, C. & Cull-Candy, S. Developmental profile of the changing properties of NMDA receptors at cerebellar mossy fiber–granule cell synapses. *J. Neurosci.* 20, 5899–5905 (2000).
- London, M. & Häusser, M. Dendritic computation. *Annu. Rev. Neurosci.* 28, 503–532 (2005).
- Kole, M. H. et al. Action potential generation requires a high sodium channel density in the axon initial segment. *Nature Neurosci.* 11, 178–186 (2008).
- Arenz, A., Silver, R. A., Schaefer, A. T. & Margrie, T. W. The contribution of single synapses to sensory representation *in vivo*. *Science* 321, 977–980 (2008).
- Markram, H. & Tsodyks, M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810 (1996).
- Abbott, L. F., Varela, J. A., Sen, K. & Nelson, S. B. Synaptic depression and cortical gain control. *Science* 275, 220–224 (1997).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This was supported by the Wellcome Trust, MRC (G0400598) and EU (EUSynapse, LSHM-CT-2005-019055). R.A.S. is in receipt of a Wellcome Senior Research Fellowship. We thank P. Gleeson for help with neuroConstruct, D. Ward and R. Kanichay for their experimental support, D. DiGregorio and P. Kirkby for discussions, and A. Arenz, D. Attwell, G. Billings, E. Chaigneau, D. DiGregorio, M. Farrant, F. Minneci and K. Vervaeke for their comments on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to R.A.S. (a.silver@ucl.ac.uk).

## METHODS

Parasagittal slices (200–250 µm) of the cerebellar vermis were prepared<sup>23,31,32</sup> from Sprague-Dawley rats (P30–40) in ACSF solution containing (in mM) 125 NaCl, 2.5 KCl, 26 NaHCO<sub>3</sub>, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 2 CaCl<sub>2</sub>, 1 MgCl<sub>2</sub>, 25 glucose and 0.5 ascorbic acid (pH 7.4 when bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>), or in a low-sodium sucrose solution containing (in mM) 85 NaCl, 2.5 KCl, 26 NaHCO<sub>3</sub>, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 0.5 CaCl<sub>2</sub>, 4 MgCl<sub>2</sub>, 25 glucose, 63.4 sucrose and 0.5 ascorbic acid. In some preparations 1 mM kynurenic acid or 10 µM D-(–)-2-amino-5-phosphonopentanoic acid (D-AP5) was added to block glutamate receptors. After ~30 min incubation at 32 °C, slices were transferred to ACSF solution at room temperature (~20 °C). Whole-cell recordings with a series resistance of 27.9 ± 1.3 MΩ ( $n = 62$ ) were made with an Axopatch 200B amplifier, filtered at 7–10 kHz, and digitized at 33–100 kHz using an InstruTech ITC-18 board and Axograph or NeuroMatic software (<http://www.neuromatic.thinkrandom.com>). GCs had a cell capacitance of 3.2 ± 0.1 pF ( $n = 62$ ) and a resting membrane potential of  $-78.5 \pm 0.9$  mV ( $n = 52$ ). Data were analysed using NeuroMatic within the Igor Pro environment (WaveMetrics).

**Synaptic conductance trains and dynamic-clamp.** Whole-cell recordings were made from GCs perfused with ACSF containing 0.3 µM strychnine and 10 µM SR95531 (Gabazine) to block inhibitory receptors, and 3 µM glycine to ensure NMDAR activation. Recordings of EPSC trains to be used for dynamic-clamp were made at –54 mV or –60 mV, using fire-polished borosilicate micropipettes containing (in mM) 110 KmeSO<sub>3</sub>, 4 NaCl, 1.78 CaCl<sub>2</sub>, 0.3 Na-GTP, 4 Mg-ATP, 40 HEPES and 5 EGTA (pH 7.3), or in some cases 90 CsCl, 10 NaCl, 1.78 CaCl<sub>2</sub>, 0.3 Na-GTP, 4 Mg-ATP, 40 HEPES, 5 EGTA and 5 TEA (pH 7.3). EPSCs were evoked by extracellular stimulation of a single MF input<sup>31</sup> using Poisson stimulation trains (PSTs) of 250–2,500 ms duration with a 1-ms minimum refractory interval. Ten repetitions of four statistically different PSTs were recorded from each MF–GC synapse at each frequency. Stimulus artefacts were removed by convolving a single stimulus artefact with the PST, and subtracting the resulting waveform from the averaged data. Synaptic current trains were converted to conductance trains by dividing by the holding potential, after correcting for the liquid junction potential<sup>13,32</sup>. Conductance trains with the same statistics from four different GCs were aligned and averaged, and had a mean conductance close to the population mean<sup>23,32</sup>. To create non-depressing AMPAR conductance trains, the first AMPAR response of each train was fit with the following multiple-exponential function:

$$G(t) = (1 - e^{-t/\tau_r})^n [a_1 e^{-t/\tau_{d1}} + a_2 e^{-t/\tau_{d2}} + a_3 e^{-t/\tau_{d3}}] \quad (1)$$

Where  $e$  is the base of the natural logarithm,  $\tau_r$  and  $\tau_{d1-3}$  are the time constants of the rising and decaying components, respectively,  $a_{1-3}$  are the amplitude components and  $t$  is time. The resulting fit was then convolved with the given PST to create a non-depressing train (Fig. 1b, red trace).

**Dynamic-clamp recordings and spike analysis.** Slices were perfused with ACSF containing 10 µM SR95531 to block tonic GABAR currents. Dynamic-clamp recordings<sup>33,34</sup> were made from GCs using fire-polished borosilicate micropipettes containing (in mM) 114 KmeSO<sub>3</sub>, 6 NaOH, 3 MgCl<sub>2</sub>, 0.02 CaCl<sub>2</sub>, 0.3 Na-GTP, 4 Na-ATP, 40 HEPES and 0.15 BAPTA (pH 7.3). A liquid junction potential of +6.3 mV ( $n = 5$ ) was corrected before gaining whole-cell access. During recordings, the resting membrane potential ( $V_{rest}$ ) was maintained near –80 mV using small amounts of holding current. Conductance trains were injected with a three-channel SM1 amplifier (Cambridge Conductance). Because GCs receive an average of four MF inputs, and often require more than one input to fire<sup>35,36</sup>, AMPAR conductance trains (and NMDAR conductance trains) from four statistically different PSTs were summated together at each input frequency ( $f$ ) before injection into GCs by means of dynamic-clamp as described below.

Unlike the AMPAR component, which follows a simple linear Ohmic relationship with voltage, the NMDAR component introduced by the SM1 amplifier has a Boltzmann-like nonlinearity that mimics the voltage-dependent Mg<sup>2+</sup> block of the NMDA conductance measured in GCs (Supplementary Fig. 4). This nonlinearity introduced by the SM1 amplifier scaled the NMDAR conductance waveforms so that the final peak value, after leaving the SM1 amplifier, matched that of GCs, as described by the following Boltzmann function:

$$G_{exc}(V) = \frac{G_{max}}{1 + e^{-(V - V_{0.5})/k}} \quad (2)$$

where  $G_{max} = 367.9$  pS,  $V_{0.5} = -12.8$  mV and  $k = 22.4$  mV ( $n = 6$  cells; curve fit to data in Supplementary Fig. 4a). During the experiments, conductance trains at various  $f$  were injected in a random order, and conditions with and without tonic inhibition ( $G_{inh} = 500$  pS; steps began 10 ms before onset of the synaptic input) were presented consecutively at each  $f$ . Some GCs were excluded because their output spike rate was too low or absent in the presence of STD and tonic inhibition (13 out of 36 cells), preventing an accurate measurement of a gain change. This will tend to underestimate the gain change we report. Cells were also

discarded if the heights of their action potentials changed more than 10% during the experiment (3 out of 36 cells). For dynamic-clamp experiments, GC output firing rate ( $F$ ) was calculated from 100 ms after the onset to the end of the stimulus train. Action potentials were detected using a threshold-level detection set at 0 mV.

**Granule cell excitation with mossy fibre stimulation.** GCs were directly activated by stimulating single MF inputs with burst PSTs ( $f = 25$ –800 Hz; 1 ms minimum refractory interval) to mimic *in vivo* MF activity<sup>19,20</sup>. Bursts lasted for a period  $T = 15/f$  seconds, consisting of approximately 15 stimuli. Average conductance and output spike rate were computed over the same period  $T$ , but starting from the time of the fourth stimulus, when STD reached approximate steady-state levels<sup>21</sup>. Recording conditions were the same as those described for the dynamic-clamp recordings, except  $V_{rest}$  was maintained at –75 mV rather than –80 mV because most GCs (69%;  $n = 22$  of 32) either did not fire action potentials when stimulated from their resting potential or did not produce enough output spikes to compute an input–output relationship in the range 25–800 Hz. Tonic inhibition was applied with dynamic-clamp ( $G_{inh} = 500$  pS; steps began 75 ms before onset of the synaptic stimulation) and was alternated with control conditions at each  $f$ .

**Data analysis.** Plots of average excitatory conductance ( $G_{exc}$ ) versus MF input rate  $f$  ( $G_{exc}f$  relationships; Figs 2b and 3b) were fit with a linear equation for AMPAR conductances without STD:

$$G_{exc}(f) = mf \quad (3)$$

and an exponential function for AMPAR conductances with STD:

$$G_{exc}(f) = m\lambda[1 - e^{-f/\lambda}] \quad (4)$$

where  $\lambda$  is a frequency constant and  $m$  is a shared slope factor. At low frequencies ( $f \ll \lambda$ ), these two equations are approximately equal. Plots of GC firing rate  $F$  versus  $G_{exc}$  ( $FG_{exc}$  relationships) were well described by a Hill equation of the form:

$$F(G_{exc}) = \frac{F_{max}}{1 + \left(\frac{G_{exc}50}{G_{exc}}\right)^n} + F_0 \quad (5)$$

where  $n$  is the exponent factor,  $F_0$  is the firing rate offset,  $F_{max}$  is the maximum firing rate and  $G_{exc}50$  is the value of  $G_{exc}$  at which  $F$  reaches half maximum. For  $n = 1$  the relationship is a simple saturating function (Fig. 2d, inset) and for  $n > 1$  the relationship is sigmoidal (Fig. 2c). Input–output relationships (Figs 1d, 3c and 4c) were also fit with equation (5), but with  $G_{exc}(f)$  substituted for  $G_{exc}$ , where  $G_{exc}(f)$  was described by equation (3) for the AMPAR component without STD, and equation (4) for STD ( $\lambda$  and  $m$  held constant during the fits). Similar results were obtained by fitting a simple Hill function (equation (5)), with  $f$  substituted for  $G_{exc}$ , but the STD input–output data were less well described by this function than when equations (4) and (5) were combined, as theoretically predicted. Fits of the input–output relationships with and without inhibition were compared using the  $F$ -ratio for the separate and combined data sets. All fits were significantly different ( $P < 0.05$ ). The gain was calculated from the average slope ( $F'$ ) of the fits between 5% and 75% its maximum value. An upper limit of 75% was used so that all computations of  $F'$  were limited to the range of our experimental data. Changes in gain ( $\Delta G$ ) were computed as follows:

$$\Delta G = \left( \frac{F'_{+a} - F'_{-a}}{F'_{-a}} \right) \quad (6)$$

where  $+a$  and  $-a$  denote conditions with and without inhibition ( $\pm inh$ ) or with and without STD ( $\pm STD$ ). Additive offset shifts ( $\Delta Offset$ ) were defined as the difference between the half-maximum frequencies of the fits for the two conditions  $+a$  and  $-a$ .

**Neuronal models.** A GC-like conductance-based integrate-and-fire model was implemented in the NEURON simulation environment<sup>37</sup> and was described by the following equation:

$$-C_m \frac{dV}{dt} = \frac{(V - E_L)}{R_m} + D(t)G_{AMPAR}(V - E_{AMPAR}) + G_{GABA}(V - E_{GABA}) \quad (7)$$

where  $C_m = 3.1$  pF, the reversal potentials for the leak  $E_L = -75$  mV, AMPARs  $E_{AMPAR} = 0$  mV and GABARs  $E_{GABA} = -75$  mV, and a membrane resistance  $R_m = 2.6$  GΩ<sup>13,32</sup>. Spikes were generated when the model reached a threshold of –49 mV, at which time the voltage was set to 10 mV for one integration time step, and then clamped to –75 mV for a refractory period of 2.5 ms. The AMPAR conductance was described by equation (1), using the following values computed from fits to our GC AMPAR conductance data:  $n = 11$ ,  $\tau_r = 0.10$  ms,  $a_1 = 2.23$  nS,  $\tau_{d1} = 0.45$  ms,  $a_2 = 0.29$  nS,  $\tau_{d2} = 2.88$  ms,  $a_3 = 0.08$  nS and  $\tau_{d3} = 21.67$  ms. For simulations with STD, whenever the integrate-and-fire model received an excitatory input, the amplitude of the postsynaptic AMPAR

response was multiplied by a scale factor  $\delta$  ( $D \rightarrow D\delta$ )<sup>38</sup>. Between inputs,  $D$  recovered exponentially back to its initial value of 1.0 with a time constant  $\tau_D = 40$  ms. For simulations with phasic inhibitory input (Supplementary Fig. 3) the synaptic GABA<sub>A</sub> conductance was described by equation (1) using the following values:  $n = 8.34$ ,  $\tau_r = 0.14$  ms,  $a_1 = 53.02$ ,  $\tau_{d1} = 0.05$  ms,  $a_2 = 0.53$ ,  $\tau_{d2} = 6.90$  ms and  $a_3 = 0$  (R. Kanichay and R.A.S., unpublished data). The peak conductance was set to 663 pS to give a time-averaged conductance of 500 pS at 100 Hz. Mean input rates were 50–150 Hz.

A model of a cortical layer 5 neuron<sup>27</sup> was synaptically connected to 400 excitatory neurons and 30 inhibitory neurons using neuroConstruct<sup>39</sup> (<http://www.NeuroConstruct.org>) and the code was automatically generated for the NEURON simulator<sup>37</sup>. Each excitatory connection cell had three synaptic contacts (to model six contacts with intermediate release probability)<sup>40</sup> and inhibitory connections had five synaptic contacts (to model ten synaptic inputs)<sup>41</sup>. Both input types were distributed randomly over the basolateral dendritic tree and soma to mimic layer 5–layer 5 connectivity. Excitatory synaptic inputs had a conductance time course described by two exponential functions ( $\tau_{rise} = 0.3$  ms,  $\tau_{decay} = 3$  ms) and a peak conductance of 3 nS<sup>40</sup> (that is, 1 nS at each of the three synaptic contacts). Each presynaptic excitatory neuron was driven to fire an independent Poisson spike train (1–70 Hz) and the synaptic conductance was scaled by the level of STD, which was implemented as described previously using experimentally measured values for layer 5 pyramidal cell connections (where the fraction of synaptic resource used,  $U_{se}$ , was equal to 0.5, and the time constants of inactivation and recovery were  $\tau_{inact} = 3$  ms and  $\tau_{recov} = 500$  ms, respectively)<sup>42</sup>. Inhibitory synaptic inputs had fixed conductances with a dual exponential time course ( $\tau_{rise} = 0.3$  ms,  $\tau_{decay} = 10$  ms) and a peak conductance of 3.5 nS<sup>41</sup>. Presynaptic inhibitory neurons were driven with independent Poisson trains with mean rates 30–150 Hz. Reversal potentials for excitatory and inhibitory conductances were 0 mV and −75 mV, respectively. Because the excitatory drive for non-depressing synapses was much higher than the more physiological depressing inputs, we carried out simulations with two different synaptic conductances (3 nS and 0.5 nS) for the non-STD case. 50 Hz inhibition produced a similar gain change for the two conductances (−11.7% and −10.9%, respectively), but the larger conductance gave a bigger gain change at 150 Hz as expected for noise-based gain modulation (−28.6% versus −22.9%). The 0.5 nS value was chosen for comparison because the excitatory input drive (0–70 Hz) better matched that of the STD case (0–60 Hz; Fig. 5b, d) than for 3 nS inputs, which saturated firing at input rates of only 12 Hz. Output firing rate was measured at steady state. Input–output relationships were fit with a simple Hill function similar to equation (5), but with input frequency ( $f$ ) substituted for  $G_{exc}$ , owing to difficulty in defining the effective  $G_{exc}$  for a distributed input.  $\Delta Gain$  and  $\Delta Offset$  were calculated from the fits as above, except for the −STD simulations, which showed little saturation of firing rate. In this case,  $\Delta Gain$  and

$\Delta Offset$  were calculated assuming a maximum rate of 100 Hz. Because the model<sup>27</sup> had an input resistance in the absence of synaptic input (11.5 MΩ) at the low end of that measured experimentally 9–123 MΩ (mean 36 MΩ)<sup>43</sup>, we repeated the simulations with a different model<sup>44</sup> with a high-input resistance (79 MΩ), which we modified to fire repetitively. We obtained similar results for the effects of inhibitory synaptic input on the input–output relationship in the presence and absence of STD in the excitatory input (data not shown), although fewer presynaptic cells (100 excitatory and 20 inhibitory) were required to drive the model over a similar range of frequencies.

31. Silver, R. A., Cull-Candy, S. G. & Takahashi, T. Non-NMDA glutamate receptor occupancy and open probability at a rat cerebellar synapse with single and multiple release sites. *J. Physiol. (Lond.)* **494**, 231–250 (1996).
32. Cathala, L., Brickle, S., Cull-Candy, S. & Farrant, M. Maturation of EPSCs and intrinsic membrane properties enhances precision at a cerebellar synapse. *J. Neurosci.* **23**, 6074–6085 (2003).
33. Robinson, H. P. & Kawai, N. Injection of digitally synthesized synaptic conductance transients to measure the integrative properties of neurons. *J. Neurosci. Methods* **49**, 157–165 (1993).
34. Sharp, A. A., O’Neil, M. B., Abbott, L. F. & Marder, E. Dynamic clamp: computer-generated conductances in real neurons. *J. Neurophysiol.* **69**, 992–995 (1993).
35. D’Angelo, E., De Filippi, G., Rossi, P. & Taglietti, V. Synaptic excitation of individual rat cerebellar granule cells *in situ*: evidence for the role of NMDA receptors. *J. Physiol. (Lond.)* **482**, 397–413 (1995).
36. Jorntell, H. & Ekerot, C. F. Properties of somatosensory synaptic integration in cerebellar granule cells *in vivo*. *J. Neurosci.* **26**, 11786–11797 (2006).
37. Hines, M. L. & Carnevale, N. T. The NEURON simulation environment. *Neural Comput.* **9**, 1179–1209 (1997).
38. Varela, J. A. et al. A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci.* **17**, 7926–7940 (1997).
39. Gleeson, P., Steuber, V. & Silver, R. A. neuroConstruct: a tool for modeling networks of neurons in 3D space. *Neuron* **54**, 219–235 (2007).
40. Markram, H., Lubke, J., Frotscher, M., Roth, A. & Sakmann, B. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol. (Lond.)* **500**, 409–440 (1997).
41. Gupta, A., Wang, Y. & Markram, H. Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* **287**, 273–278 (2000).
42. Tsodyks, M. V. & Markram, H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA* **94**, 719–723 (1997); erratum **94**, 5495 (1997).
43. Schwindt, P. C., O’Brien, J. A. & Crill, W. E. Quantitative analysis of firing properties of pyramidal neurons from layer 5 of rat sensorimotor cortex. *J. Neurophysiol.* **77**, 2484–2498 (1997).
44. Mainen, Z. F., Joerges, J., Huguenard, J. R. & Sejnowski, T. J. A model of spike initiation in neocortical pyramidal neurons. *Neuron* **15**, 1427–1439 (1995).

## LETTERS

# The Fas–FADD death domain complex structure unravels signalling by receptor clustering

Fiona L. Scott<sup>1†</sup>, Boguslaw Stec<sup>1</sup>, Cristina Pop<sup>1</sup>, Małgorzata K. Dobaczewska<sup>1</sup>, JeongEun J. Lee<sup>1</sup>, Edward Monosov<sup>1</sup>, Howard Robinson<sup>2</sup>, Guy S. Salvesen<sup>1</sup>, Robert Schwarzenbacher<sup>3</sup> & Stefan J. Riedl<sup>1</sup>

The death inducing signalling complex (DISC) formed by Fas receptor, FADD (Fas-associated death domain protein) and caspase 8 is a pivotal trigger of apoptosis<sup>1–3</sup>. The Fas–FADD DISC represents a receptor platform, which once assembled initiates the induction of programmed cell death. A highly oligomeric network of homotypic protein interactions comprised of the death domains of Fas and FADD is at the centre of DISC formation<sup>4,5</sup>. Thus, characterizing the mechanistic basis for the Fas–FADD interaction is crucial for understanding DISC signalling but has remained unclear largely because of a lack of structural data. We have successfully formed and isolated the human Fas–FADD death domain complex and report the 2.7 Å crystal structure. The complex shows a tetrameric arrangement of four FADD death domains bound to four Fas death domains. We show that an opening of the Fas death domain exposes the FADD binding site and simultaneously generates a Fas–Fas bridge. The result is a regulatory Fas–FADD complex bridge governed by weak protein–protein interactions revealing a model where the complex itself functions as a mechanistic switch. This switch prevents accidental DISC assembly, yet allows for highly processive DISC formation and clustering upon a sufficient stimulus. In addition to depicting a previously unknown mode of death domain interactions, these results further uncover a mechanism for receptor signalling solely by oligomerization and clustering events.

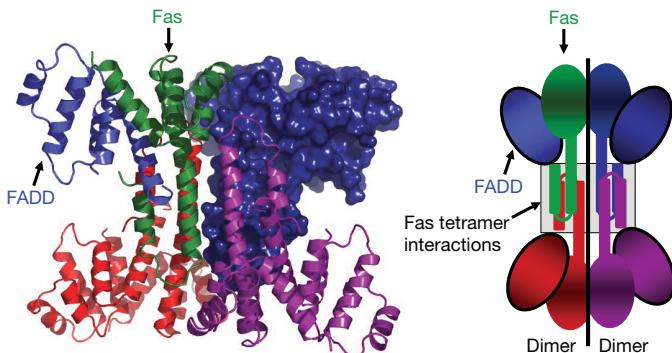
As with other oligomeric signalling platforms, the DISC acts as a cellular switch, which exists in the ‘off’ position in the absence of a stimulus followed by oligomerization of its constituents to form the active (‘on’ position) oligomeric platform<sup>6</sup>. In the case of the Fas–FADD DISC the apoptotic signal per se is the binding of Fas ligand (FasL), which in a basic view (Supplementary Fig. 1a) leads to recruitment of FADD via death domain interactions. FADD in turn recruits caspase 8 through death effector domain (DED) interactions leading to activation of this apical caspase<sup>1,2,7</sup>. However, from a cell signalling point of view this string of events remains on a descriptive level, as the binding partners do not (and should not) interact in the absence of a sufficient stimulus. Productive DISC formation is only observed when ligand binding occurs in a permissive environment, such as predisposition of Fas in membrane rafts, and is furthermore characterized by the formation of highly oligomeric DISC clusters<sup>8–14</sup>. Because the death domain interaction of Fas and FADD is at the heart of the DISC interaction network (Supplementary Fig. 1b), the question arises about how a simple death domain interaction mediates such complexity. Other signalling platforms, for example the apoptosisome or inflammasomes<sup>15,16</sup>, are proposed to elegantly exploit a nucleotide-dependent regulation to properly signal oligomerization as opposed to the seemingly simple death domain interaction of Fas and FADD.

In addition to mutations implicated in disease states, a variety of mutants have been generated in order to define interfaces in the Fas–FADD complex (Supplementary Table 2; reviewed in refs 5, 14). However, despite this information and effort, the nature of the primary Fas–FADD interaction and its implication for the actual mechanism of formation of the Fas–FADD death domain network remains elusive<sup>14</sup>. In our study we were able to gain an insight into this mechanism by elucidating the crystal structure of the Fas–FADD death domain complex from which we conclude a model explaining how observed contacts are mediated by a primary Fas–FADD complex, which acts as a sensitive switch governing DISC formation. This extends and completes our view of death domains, which initiated from rigid defined domain complexes such as the constitutive 1:1 complex seen in the APAF-1–caspase 9 CARD–CARD interaction, to current views which include plasticity and asymmetry of complexes in the death domain superfamily<sup>5</sup> (see also Supplementary Discussion). Thus, the model presented here can furthermore serve as a template for other signalling platforms lacking enzymatic components which are solely mediated by oligomerization and clustering events.

Despite its elusiveness and solubility difficulties<sup>14</sup>, we succeeded in producing a soluble Fas–FADD complex when *Escherichia coli* lysates of recombinantly expressed Fas death domain and FADD death domain were combined before purification. After purification we optimized conditions to obtain crystals of the complex and solved the 2.7 Å resolution structure. Crystals indexed in the hexagonal space group *P*6<sub>1</sub> with two tetrameric assemblies each comprising four Fas death domains and four FADD death domains in the asymmetric unit (Fig. 1 and Supplementary Table 1). The tetrameric arrangement represents a dimer of two Fas–FADD death domain complex dimers. In this arrangement Fas provides all contacts (see also Supplementary Discussion). All residues of the Fas death domain are well defined in the electron density. FADD is defined from G93 to G191 with the region around residues 107–149 showing more diffuse density due to a high degree of overall motion caused by a lack of crystal contacts (Supplementary Fig. 1c).

Inspection of the interfaces shows that the actual Fas–FADD complex is formed by helix one and six of the FADD death domain through employment of a hydrophobic patch surrounded by polar residues (Fig. 2a, d). Unlike the FADD death domain, the Fas death domain in the complex deviates from the typical death domain fold<sup>17–19</sup>. Comparison with the solution structure of the isolated Fas death domain<sup>20</sup> reveals that Fas in the complex has undergone an opening in which helix six has shifted and fused with helix five to form a long helix that we term the stem helix (Fig. 2b, c). In addition to Fas opening, a new helix at the carboxy terminus of Fas, dubbed the C-helix, is observed (Fig. 2a). The main interaction interface on

<sup>1</sup>Program in Apoptosis and Cell Death Research, The Burnham Institute for Medical Research, La Jolla, California 92037, USA. <sup>2</sup>Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, USA. <sup>3</sup>Department of Molecular Biology, University of Salzburg, 5020 Salzburg, Austria. <sup>†</sup>Present address: Apoptos Inc., San Diego, California 92121, USA.



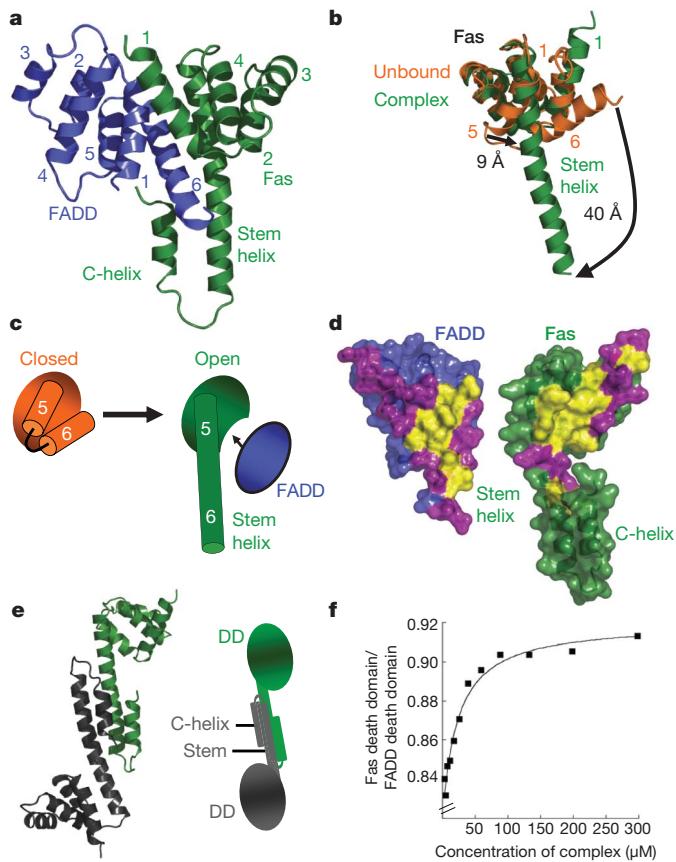
**Figure 1 | Overall structure of the Fas-FADD death domain complex.** The structure (left panel) shows a tetrameric arrangement of Fas–FADD death domain complexes. Contacts between the complexes are solely mediated by Fas molecules. The right panel shows a cartoon representation of the Fas–FADD complex structure. One Fas–FADD death domain complex is displayed in green (Fas) and blue (FADD) ribbons, whereas the remaining three complexes are coloured red, magenta and blue (surface representation).

Fas is formed by hydrophobic residues that become available only upon Fas opening (Fig. 2d). Thus, the opening of Fas discovered in the crystal structure is a crucial event in its interaction with FADD. Notably, other FADD interacting death receptors show similar residues as those observed in Fas at the complex interface (see Supplementary Fig. 4c and Supplementary Discussion).

Broader investigation of the Fas–FADD interface divulges the disperse nature of the interaction surface, which lacks the presence of defined and focused interaction sites often referred to as hotspots<sup>21–23</sup>. This phenomena is used in regulatory complexes characterized by weak binding despite sufficiently large interaction surfaces<sup>22,24,25</sup>. In addition to the lack of hotspots, a loss of compact tertiary structure owing to the opening of Fas further enhances flexibility of the system and thus additionally weakens the complex. Besides providing the primary interface for binding of FADD, a second consequence of the opening of Fas is to allow for interaction with another open Fas molecule forming a Fas–Fas bridge. The predominant Fas–Fas association observed in the structure is formed by stem helices and C-helices of the two Fas molecules building a Fas–Fas dimer (Fig. 2e), which compared to the tetramer, better approximates structural requirements in the context of the natural membrane association (see Supplementary Discussion). Thus, opening of Fas provides both the basis for FADD binding and the formation of a regulatory bridge between Fas molecules. Owing to the difficult nature of the Fas–FADD complex, biochemical analysis is limited (Supplementary Fig. 2a), but indeed confirms a weak primary Fas–FADD complex which dissociates at concentrations as high as 50 μM (Fig. 2f and Supplementary Fig. 2d). Mutational analysis of the Fas–FADD interface reveals predominantly mild phenotypes (Supplementary Fig. 3a–c, and data not shown), in line with previous observations for complexes lacking binding hotspots such as IFNγ2/IFNAR1<sup>26</sup>.

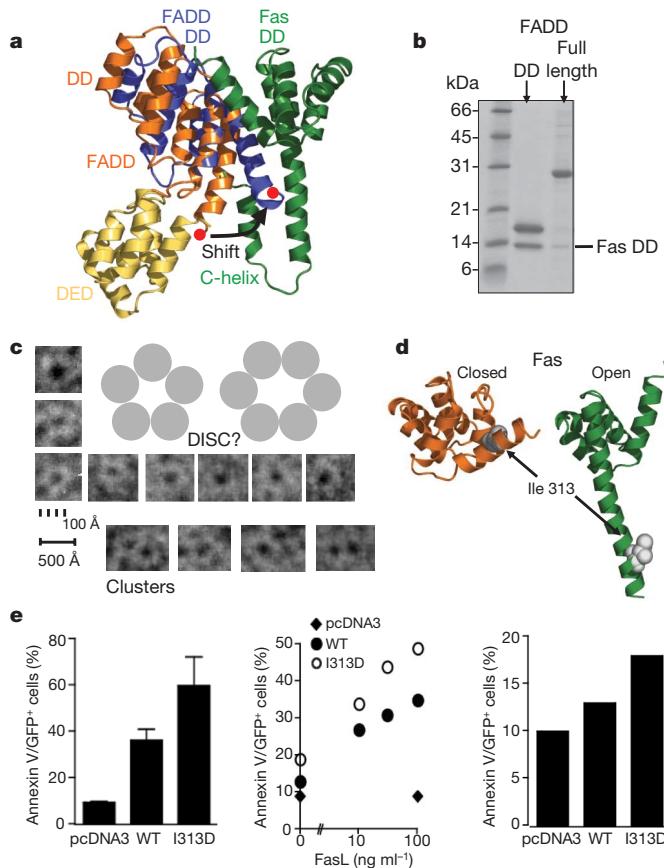
In solution, the Fas–FADD death domain complex indeed shows a tetrameric behaviour (Supplementary Fig. 2b–d), which is further corroborated by previously reported results of a solubility enhanced complex that also pointed to a tetrameric arrangement<sup>27</sup>. Taken together, these results establish that the observed tetramer in the crystal structure reflects the Fas–FADD death domain arrangement present in solution and is not a result of crystal formation. In addition, the tendency of the tetramer to form higher oligomers can be observed (Supplementary Fig. 2c, d).

To obtain further insight into the binding of FADD, we overlaid the structure of full-length FADD<sup>18</sup> onto the death domain of FADD in the primary complex (Fig. 3a). The marked conformational



**Figure 2 | Fas-FADD death domain complex: Fas-FADD and Fas-Fas interactions are dependent on Fas opening.** **a**, Primary Fas–FADD death domain complex. FADD (blue) adopts the characteristic death domain fold with helices one and six participating in the main interaction site. In Fas (green) only helices one to four approximately adopt a death-domain-like fold, whereas a long helix (stem helix) is found in place of helices five and six, which together with helix one provides the main interaction residues for FADD binding. Additionally, a helix at the C terminus of Fas (C-helix) is observed. **b**, Conformational change in Fas. Comparison of the structure of unbound Fas death domain (closed form in orange; Protein Data Bank entry 1DDF) and Fas in the Fas–FADD complex. Owing to formation of the stem helix, residues of helix five and six shift significantly. Furthermore, the rearrangement of helix six exposes part of the hydrophobic core of Fas. **c**, Cartoon illustration of Fas opening. **d**, Primary Fas–FADD interface. View onto interfaces governing primary complex formation. Surface representation shows complementary hydrophobic patches (yellow) on FADD and Fas surrounded by polar residues (magenta). The hydrophobic interface on Fas becomes exposed upon Fas opening. **e**, Fas dimer unit. Another consequence of Fas opening is the formation of Fas dimer units, which interact via the stem and C-helix. The right panel is a cartoon representation of the dimer. DD, death domain. **f**, The Fas–FADD death domain complex is weak. A dilution experiment of the isolated Fas–FADD complex shows cooperative dissociation of the complex below concentrations as high as 50 μM (plot derived from quantitative SDS–PAGE analysis of Fas death domain retained on Ni-NTA resin from various Fas death domain/FADD death domain-His<sub>6</sub> complex dilutions).

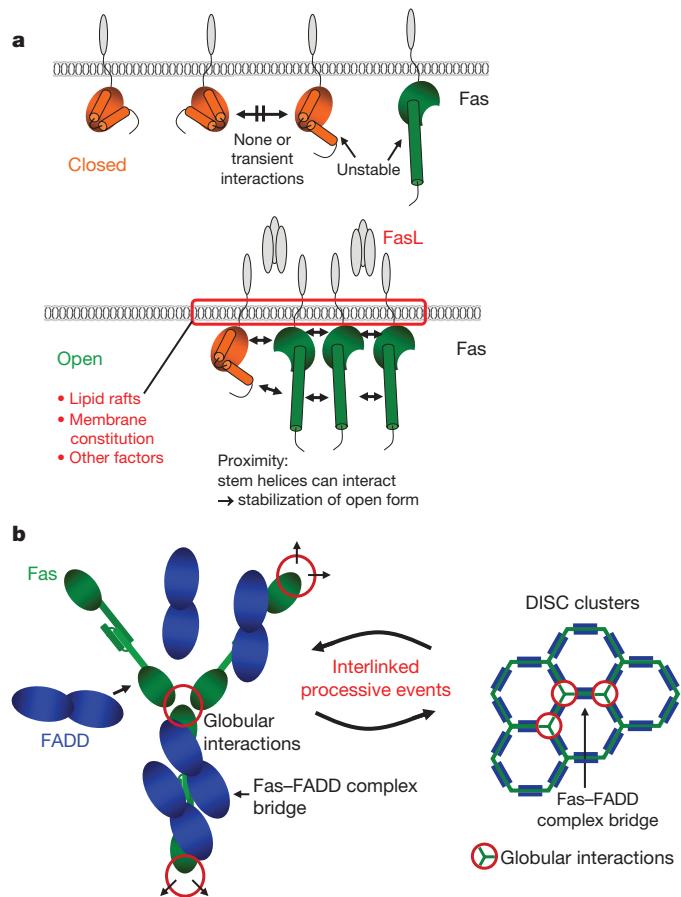
change in Fas, which is essential to bind FADD, is also accompanied by a change in the FADD death domain. The C-terminal helix of FADD has to shift to avoid a steric clash with the newly formed C-helix of Fas, substantiating results that attribute a negative role in cell death induction to this region of Fas (Fig. 3a and Supplementary Fig. 3d)<sup>28</sup>. This observation is even more crucial in the context of full-length FADD, as the FADD C terminus is located at the interface of the DED and death domain. Along this line, an overall conformational rearrangement, namely the relative position of the DED and the death domain of FADD, can also be concluded to expose caspase-8-binding residues determined in earlier studies



**Figure 3 | The Fas-FADD bridge in the DISC: binding of full-length FADD and the key role of Fas opening *in vivo*.** **a**, Overlay of the structure of full-length FADD (Protein Data Bank entry 2gf5; orange) onto the Fas-FADD complex structure. (Full-length FADD refers to the well characterized FADD F25Y mutant; see Supplementary Methods.) The last helix of the death domain of unbound FADD (red dot) shifts to avoid clashing with the newly formed Fas C-helix in the Fas-FADD complex. **b**, Conformational change in full-length FADD. Proteins were expressed separately, and His-tagged versions of the death domain of FADD, or full-length FADD, were mixed with untagged Fas. Ni-NTA chromatography demonstrates that full-length FADD shows reduced initial binding to the Fas death domain, when compared to the FADD death domain protein. **c**, Although initial binding of full-length FADD to Fas death domain is reduced, prolonged incubation leads to the formation of DISC-like structures. Incubation of both proteins overnight led to the formation of ring-like structures with a strong tendency to form clusters as determined using electron microscopy. Displayed are single ring-like structures and clusters from several images. Owing to their strong tendency to self-adhere, no consistent monolayer for thorough evaluation could be generated so far. **d**, Location of Ile 313 in closed (unbound, orange) and open (complex, green) forms of Fas. **e**, Huh7 cells transfected with Fas I313D show elevated cell death, as assessed by annexin V reactivity, compared with Fas wild type on stimulation with Fas antibody (left, error bars indicate standard error of the mean,  $n = 3$ ), FasL (middle), and also in the absence of a stimulus (right). Equal cell surface Fas expression was confirmed by FACS analysis and immunoblot (data not shown).

(summarized in refs 5, 18). This change is readily enabled by the interdomain linker, which is composed exclusively of Ala, Gly and Pro residues. Indeed, our studies found that when full-length FADD is incubated with Fas a reduced initial complex formation can be observed, underlining a conformational change (Fig. 3b), whereas longer incubation of the two proteins leads to the formation of protein clusters (Fig. 3c).

Details revealed in the Fas-FADD death domain structure show that Fas opening is key to both FADD binding and formation of a Fas-Fas bridge. To validate the relevance of Fas opening *in vivo*, we sought to create a mutation that interfered with the closed form of



**Figure 4 | Model of DISC formation: mechanism of receptor signalling through clustering.** Schematic diagram illustrating DISC formation. **a**, Model for Fas opening. Like in any two-state model, it can be assumed that the closed and open forms of Fas exist in equilibrium. In the absence of an apoptotic signal the equilibrium between closed and open forms of Fas strongly favours the closed form whereas open forms are unstable. Upon an apoptotic stimulus the equilibrium shifts to favour the open form of Fas. Fas molecules are brought together by Fas ligand in a permissive environment dependent on several factors, including lipid rafts and membrane constitution. The close proximity of Fas death domains now allows for stem helices to interact, leading to stabilization of the open form. **b**, Opening and formation of the Fas-Fas bridge links trimeric death domain units defined in earlier studies (Supplementary Table 2), which are formed by globular regions of the open Fas. FADD can now bind to the open Fas molecules, further stabilizing the bridge. The consequence is processive interlinked DISC formation and clustering in which open Fas molecules interacting via their globular domains are linked by a multitude of weak Fas-FADD bridges, leading to overall stable DISC clusters. This permits activation of caspase 8, presumably by proximity-enforced dimerization<sup>30</sup>, and induction of apoptosis.

Fas and propagated Fas opening. If Fas opening is a prerequisite for DISC formation, this forced opening of Fas should produce a hyperactive phenotype. By comparing the structures of closed and open Fas we identified Ile 313 as a candidate. Ile 313 is part of the hydrophobic core and resides in helix 6 of the closed form and is thus at the centre of the opening event. At the same time it does not directly participate in the Fas-FADD interaction or the observed Fas-Fas dimer in the complex structure (Fig. 3d and Supplementary Fig. 4a). Indeed, mutation of Ile 313 to Asp (I313D) resulted in significantly enhanced killing. Enhanced cell death by Fas I313D compared to wild type could be observed when cells were stimulated with Fas antibody or FasL and even in the absence of a stimulus, further establishing the central role of Fas opening (Fig. 3e).

On the basis of the analysis of the Fas-FADD structure, which displays key regulatory features of DISC formation, and taking into

account previous findings, we conclude a model of how receptor signalling in the DISC occurs solely through clustering events. In this model the opening of Fas is central to form the Fas–Fas bridge and to recruit FADD. The Fas–Fas bridge and FADD binding are governed by weak protein–protein interactions only stabilized upon processive clustering, thus representing the actual regulatory switch governing DISC formation (Fig. 4).

Until now, the dominant model for DISC formation proposed that trimeric Fas ligand crosslinked units of preformed trimeric Fas in local membrane rafts<sup>8–14</sup>, leading to higher order arrangements. This model is substantiated by mutations in the Fas death domain (Supplementary Table 2; reviewed in refs 5, 14) that are not located at the primary Fas–FADD interface, but rather adjacently on the globular units of open Fas observed here. While illustrating the overall process, this initial model did not provide a mechanism for the actual clustering of Fas death domains within the cell and most importantly the relationship between Fas clustering and FADD binding, both of which are explained in our model. In this model only when a sufficient number of Fas molecules are in close proximity, as is expected on Fas ligand binding, can the open forms of Fas become stabilized (Fig. 4a), because the newly formed stem helices can interact with each other. An interlinked consequence of this interaction is that Fas molecules become connected through this sensitive bridge—setting the stage for a chain reaction. Now the formation of Fas bridges linking the Fas death domain trimeric units—which interact via the globular portion of open Fas—directly results in rapid and processive clustering on the inside of the cell membrane (Fig. 4b). Also, FADD is now able to bind the open Fas molecules, key for DISC signalling, and additionally increasing stabilization of the Fas–Fas bridges, thereby further fostering DISC formation and clustering (Fig. 4b).

The proximity-induced stabilization of open Fas, in combination with the formation of the weak bridge, allows timing and strict regulation of DISC formation solely by clustering as opposed to, for example, a nucleotide-exchange-mediated oligomerization as in the apoptosome<sup>29</sup>.

## METHODS SUMMARY

The Fas–FADD death domain complex was obtained by mixing crude *E. coli* lysates from recombinantly expressed proteins at high concentration. Large amounts of pure, soluble and crystallization-grade complex could be obtained after Ni-affinity and ion-exchange chromatography. Crystallization could be achieved under acidic conditions, which after several rounds of optimization resulted in diffracting crystals, one to 2.7 Å resolution. Initial phases were obtained using a Pt-derivative crystal followed by standard structure solution using the native data.

Received 27 August; accepted 30 October 2008.

Published online 31 December 2008.

- Ashkenazi, A. & Dixit, V. M. Apoptosis control by death and decoy receptors. *Curr. Opin. Cell Biol.* **11**, 255–260 (1999).
- Peter, M. E. & Krammer, P. H. The CD95(APO-1/Fas) DISC and beyond. *Cell Death Differ.* **10**, 26–35 (2003).
- Taylor, R. C., Cullen, S. P. & Martin, S. J. Apoptosis: controlled demolition at the cellular level. *Nature Rev. Mol. Cell Biol.* **9**, 231–241 (2008).
- Fesik, S. W. Insights into programmed cell death through structural biology. *Cell* **103**, 273–282 (2000).
- Park, H. H. et al. The death domain superfamily in intracellular signaling of apoptosis and inflammation. *Annu. Rev. Immunol.* **25**, 561–586 (2007).
- Festjens, N., Cornelis, S., Lamkanfi, M. & Vandenebeeck, P. Caspase-containing complexes in the regulation of cell death and inflammation. *Biol. Chem.* **387**, 1005–1016 (2006).
- Tibbetts, M. D., Zheng, L. & Lenardo, M. J. The death effector domain protein family: regulators of cellular homeostasis. *Nature Immunol.* **4**, 404–409 (2003).
- Algeciras-Schimmler, A. et al. Molecular ordering of the initial signaling events of CD95. *Mol. Cell. Biol.* **22**, 207–220 (2002).
- Feig, C., Tchikov, V., Schutze, S. & Peter, M. E. Palmitoylation of CD95 facilitates formation of SDS-stable receptor aggregates that initiate apoptosis signaling. *EMBO J.* **26**, 221–231 (2007).

- Lavrik, I. N. et al. Analysis of CD95 threshold signaling: triggering of CD95 (FAS/APO-1) at low concentrations primarily results in survival signaling. *J. Biol. Chem.* **282**, 13664–13671 (2007).
- Muppudi, J. R. & Siegel, R. M. Ligand-independent redistribution of Fas (CD95) into lipid rafts mediates clonotypic T cell death. *Nature Immunol.* **5**, 182–189 (2004).
- O'Reilly, L. A. et al. Modifications and intracellular trafficking of FADD/MORT1 and caspase-8 after stimulation of T lymphocytes. *Cell Death Differ.* **11**, 724–736 (2004).
- Siegel, R. M. et al. SPOTS: signaling protein oligomeric transduction structures are early mediators of death receptor-induced apoptosis at the plasma membrane. *J. Cell Biol.* **167**, 735–744 (2004).
- Werner, M. H., Wu, C. & Walsh, C. M. Emerging roles for the death adaptor FADD in death receptor avidity and cell cycle regulation. *Cell Cycle* **5**, 2332–2338 (2006).
- Petrilli, V., Dostert, C., Muruve, D. A. & Tschoop, J. The inflammasome: a danger sensing complex triggering innate immunity. *Curr. Opin. Immunol.* **19**, 615–622 (2007).
- Riedl, S. J. & Salvesen, G. S. The apoptosome: signalling platform of cell death. *Nature Rev. Mol. Cell Biol.* **8**, 405–413 (2007).
- Berglund, H. et al. The three-dimensional solution structure and dynamic properties of the human FADD death domain. *J. Mol. Biol.* **302**, 171–188 (2000).
- Carrington, P. E. et al. The structure of FADD and its mode of interaction with pro-caspase-8. *Mol. Cell* **22**, 599–610 (2006).
- Jeong, E. J. et al. The solution structure of FADD death domain. Structural basis of death domain interactions of Fas and FADD. *J. Biol. Chem.* **274**, 16337–16342 (1999).
- Huang, B., Eberstadt, M., Olejniczak, E. T., Meadows, R. P. & Fesik, S. W. NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain. *Nature* **384**, 638–641 (1996).
- Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383–386 (1995).
- Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H. & Schreiber, G. The molecular architecture of protein-protein binding sites. *Curr. Opin. Struct. Biol.* **17**, 67–76 (2007).
- Bogdan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1–9 (1998).
- Huxford, T. et al. Solvent exposed non-contacting amino acids play a critical role in NF-κB/IκBα complex formation. *J. Mol. Biol.* **324**, 587–597 (2002).
- Janin, J., Miller, S. & Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155–164 (1988).
- Roisman, L. C., Jaitin, D. A., Baker, D. P. & Schreiber, G. Mutational analysis of the IFNAR1 binding site on IFNα2 reveals the architecture of a weak ligand-receptor binding site. *J. Mol. Biol.* **353**, 271–281 (2005).
- Ferguson, B. J. et al. Biophysical and cell-based evidence for differential interactions between the death domains of CD95/Fas and FADD. *Cell Death Differ.* **14**, 1717–1719 (2007).
- Itoh, N. & Nagata, S. A novel protein domain required for apoptosis. Mutational analysis of human Fas antigen. *J. Biol. Chem.* **268**, 10932–10937 (1993).
- Kim, H. E., Du, F., Fang, M. & Wang, X. Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1. *Proc. Natl. Acad. Sci. USA* **102**, 17545–17550 (2005).
- Boatright, K. M. et al. A unified model for apical caspase activation. *Mol. Cell* **11**, 529–541 (2003).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. Snipes for protein sequencing and technical assistance, J. Reed for providing Fas cDNA and A. Bobkov for the AUC. This work was supported by a P30 CA030199 cancer center grant and R01AA017238 to S.J.R.; PO1CA69381 to G.S.S.; and MCEXT-033534 to R.S. Data measured at beamline X29 of the National Synchrotron Light Source were also supported by Biological and Environmental Research DOE, and National Center for Research Resources NIH. Earlier stages of the work were supported by a LLS scholarship to S.J.R. S.J.R. is currently a V Foundation scholar.

**Author Contributions** F.L.S. performed and evaluated *in vivo* studies. J.J.L. and M.K.D. performed cloning, protein expression and crystallization. C.P. performed biochemical analyses. H.R., B.S. and in particular R.S. performed data collection and structure solution. E.M. performed EM-studies. G.S.S. participated in study design and evaluation. All authors discussed the work. S.J.R. participated in and oversaw all aspects of the work and wrote the manuscript.

**Author Information** Atomic coordinates and experimental structure factors have been deposited within the Protein Data Bank and are accessible under the code 3EZQ. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.J.R. (sriedl@burnham.org) or R.S. (robert.schwarzenbacher@sbg.ac.at).

## LETTERS

# Artificial nanopores that mimic the transport selectivity of the nuclear pore complex

Tijana Jovanovic-Talisman<sup>1</sup>, Jaclyn Tetenbaum-Novatt<sup>2</sup>, Anna Sophia McKenney<sup>2</sup>, Anton Zilman<sup>3</sup>, Reiner Peters<sup>4</sup>, Michael P. Rout<sup>2</sup> & Brian T. Chait<sup>1</sup>

Nuclear pore complexes (NPCs) act as effective and robust gateways between the nucleus and the cytoplasm, selecting for the passage of particular macromolecules across the nuclear envelope. NPCs comprise an elaborate scaffold that defines a ~30 nm diameter passageway connecting the nucleus and the cytoplasm. This scaffold anchors proteins termed ‘phenylalanine-glycine’ (FG)-nucleoporins, the natively disordered domains of which line the passageway and extend into its lumen<sup>1</sup>. Passive diffusion through this lined passageway is hindered in a size-dependent manner. However, transport factors and their cargo-bound complexes overcome this restriction by transient binding to the FG-nucleoporins<sup>2–10</sup>. To test whether a simple passageway and a lining of transport-factor-binding FG-nucleoporins are sufficient for selective transport, we designed a functionalized membrane that incorporates just these two elements. Here we demonstrate that this membrane functions as a nanoselective filter, efficiently passing transport factors and transport-factor–cargo complexes that specifically bind FG-nucleoporins, while significantly inhibiting the passage of proteins that do not. This inhibition is greatly enhanced when transport factor is present. Determinants of selectivity include the passageway diameter, the length of the nanopore region coated with FG-nucleoporins, the binding strength to FG-nucleoporins, and the antagonistic effect of transport factors on the passage of proteins that do not specifically bind FG-nucleoporins. We show that this artificial system faithfully reproduces key features of trafficking through the NPC, including transport-factor-mediated cargo import.

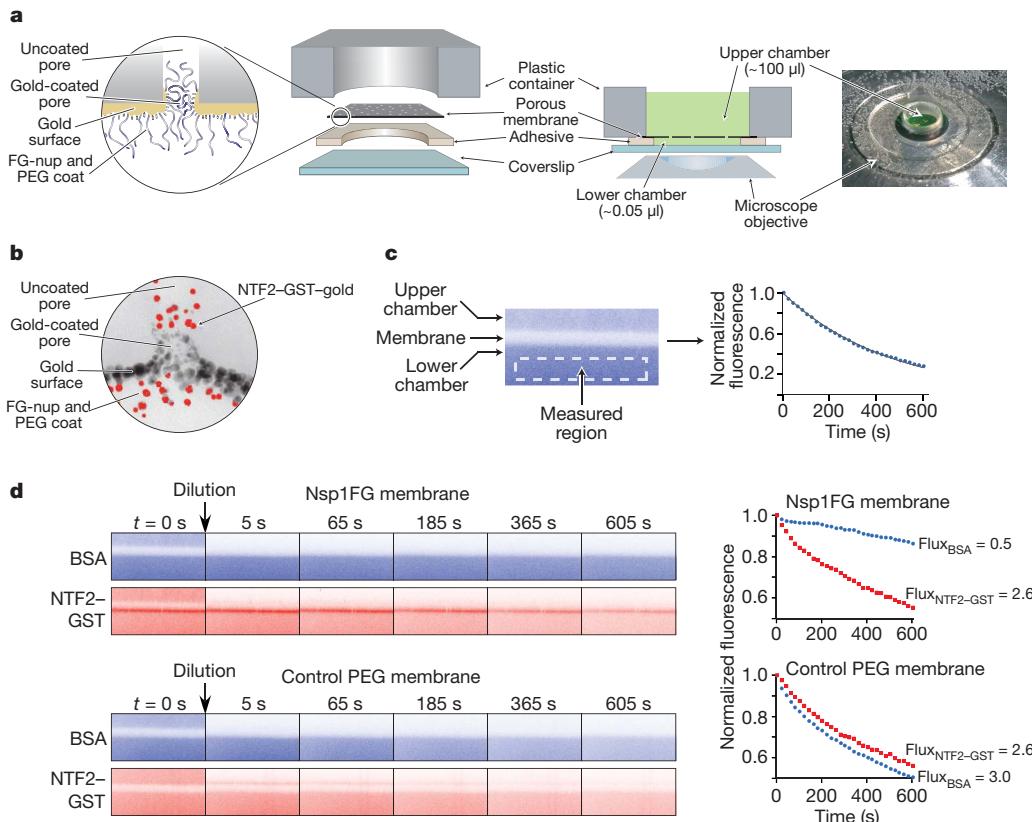
Several groups have used functionalized nanoporous membranes to selectively enrich between molecules such as chiral enantiomers<sup>11,12</sup>, compounds of differing hydrophilicities<sup>13</sup> or single-base mismatched oligonucleotides<sup>14</sup>, or more recently to selectively transport single-stranded DNA bound to polymer carrier molecules<sup>15</sup>. Similarly, the central element of our artificial system is a polycarbonate membrane perforated by ~30-nm-diameter cylindrical nanopores (Fig. 1a). The membrane is coated on one face with a thin layer of gold, to which is conjugated a single layer of a natively disordered FG-repeat domain from an FG-nucleoporin; these domains are characterized by the presence of multiple repeats containing degenerate ‘Phe-Gly’ (FG) motifs separated by low complexity hydrophilic spacers of approximately 5–30 amino acids in length. FG-repeat domains of similar size were selected from either of the nucleoporins Nsp1 or Nup100 from budding yeast. These two domains, referred to as Nsp1FG and Nup1FG, were chosen because they represent two major classes of FG-repeat motif composition (FxFG- and GLFG-repeats, respectively)<sup>16</sup> (Supplementary Fig. 1). Our functionalized membrane is designed to mimic the following essential properties of

NPCs: it has pores of similar diameter to those in NPCs; each pore is coated by a monolayer of FG-nucleoporins; these FG-nucleoporins are properly oriented, surrounding and partially occluding the pore opening as in the NPC; they are at approximately the same density as found in the NPC (~70 molecules per pore, Supplementary Methods); and transport takes place across a thin (~15 nm) barrier (Fig. 1a), akin to the nuclear envelope, with the transported material entering and (crucially) exiting the NPC mimic.

We explored the transport properties of our functionalized membranes by either changing the material coating the membrane and pores or changing the proteins passing across the pores. The flux of fluorescently labelled proteins through the functionalized membranes was measured using the device illustrated in Fig. 1a, a setup inspired by that used for optical single transporter recording of NPCs<sup>17</sup>. In each case, the proteins to be measured were allowed to reach equilibrium between the small lower chamber and the much larger upper chamber. The solution in the upper chamber was then rapidly diluted ~50-fold and the flux of protein diffusing from the lower to upper chamber was measured by quantifying the decrease in fluorescent material in the lower chamber (Fig. 1c). Before investigating the properties of the FG-nucleoporin functionalized membranes, we measured protein fluxes through a control membrane, coated with a layer of an inert low molecular weight (356 daltons (Da)) compound (small PEG). As expected, the resulting fluxes were largely dependent on protein size (Supplementary Fig. 5 and Supplementary Tables 2–5).

We then measured the flux of proteins across membranes when both a transport factor and a control protein were present at the same time. In each case, the transport factor can specifically bind the FG repeats with relatively high affinity, whereas the control can at most form weak, nonspecific interactions (Supplementary Fig. 10c, d). Each protein was labelled using a different fluorophore (that is, two-colour experiment). We measured the flux of the dimeric transport factor human nuclear transport factor 2-glutathione S-transferase (NTF2-GST) (82 kDa) versus the control bovine serum albumin (BSA; a protein of similar mass (66 kDa), Stokes radius and isoelectric point to the NTF2-GST dimer) through Nsp1FG-coated membranes (Fig. 1d). The flux of BSA across the Nsp1FG membrane was on average threefold lower than that of NTF2-GST (Supplementary Table 3). In contrast, in an assembly containing a small-PEG-coated membrane, the same mixture of NTF2-GST and BSA showed equally rapid efflux of both proteins out of the lower chamber, as expected from two such comparably sized proteins through essentially open holes (Fig. 1d). Thus, the Nsp1FG-coated pores were as permeable to NTF2-GST as the small-PEG-coated pores, but simultaneously much less permeable to BSA. Figure 2a shows ‘flux

<sup>1</sup>Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, and <sup>2</sup>Laboratory of Cellular and Structural Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA. <sup>3</sup>Theoretical Biology and Biophysics Group and Center for Nonlinear Studies, Theoretical Division, Los Alamos National Laboratory, PO Box 1663, Los Alamos, New Mexico 87545, USA. <sup>4</sup>Institute of Medical Physics and Biophysics and Center of Nanotechnology (CeNTech), University of Muenster, Robert-Koch-Strasse 31, 48149 Muenster, Germany.



**Figure 1 | Design and operation of the NPC mimic.** **a**, Far left: schematic of a single pore in the functionalized membrane. A 6- $\mu$ m-thick polycarbonate membrane perforated by  $\sim$ 30-nm channels ( $8 \times 10^8$  pores per  $\text{cm}^2$ ) was coated on one face with a  $\sim$ 15-nm-thick gold layer to which we attached FG-nucleoporins (FG-nups) by a single carboxy-terminal cysteine; hence, all FG-nucleoporins were similarly oriented. Subsequent attachment of 356-Da PEG-thiol (small PEG) molecules blocked any remaining exposed gold surface<sup>8,9</sup>. Centre left: exploded view of the device carrying the membrane. Centre right: sectional view of the device mounted on a confocal microscope and loaded with a fluorescently labelled protein solution (green). Lower chamber: height,  $\sim$ 25  $\mu$ m; diameter,  $\sim$ 1.6 mm. Upper chamber: height,  $\sim$ 1 mm; diameter,  $\sim$ 2 mm; overflow capacity, 100  $\mu$ l. Far right: top-view photograph of the device. **b**, Transmission electron micrograph of a single pore on Nsp1FG-functionalized membrane, incubated with gold-labelled NTF2-GST (pseudocoloured red). NTF2-GST is seen to bind the FG-nucleoporin layer projecting from the gold surface and transit into the 30-nm pore. **c**, Method for measurement of protein fluxes across the

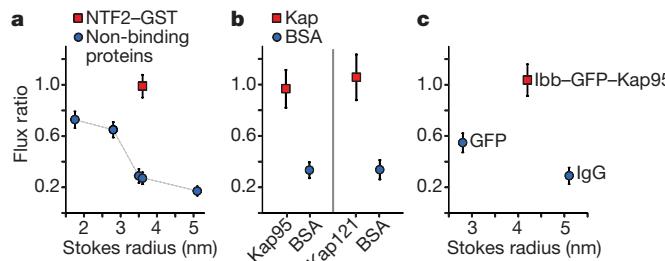
nano-selective filter. Left: confocal microscopy z-axis section through the lower chamber, membrane and upper chamber showing the fluorescence signal (blue) from a single protein in equilibrium between the two chambers. The measured volume is indicated by the dashed box. Right: plot showing the decrease in fluorescence signal in the lower chamber after dilution of the upper chamber, resulting from an efflux of protein from the lower to upper chamber (Supplementary Information). **d**, Two-channel fluorescence measurements of simultaneous diffusion of fluorescein isothiocyanate (FITC)-labelled BSA (blue) and Cy5-labelled NTF2-GST (red) through either an Nsp1FG-coated membrane (top) or a control small-PEG-coated membrane (bottom). Left: time course of confocal images collected as in **c**. Colours were altered for clarity; the same data with unmodified colours are provided in Supplementary Fig. 12. Right: corresponding fluorescence decrease curves and fluxes (in molecules pore $^{-1}$  s $^{-1}$   $\mu\text{M}^{-1}$ ). The flux of NTF2-GST was similar through both membranes, whereas the flux of BSA was significantly lower through just the Nsp1FG-coated membrane.

ratios' (a measure of transport efficiency) for several control proteins of various sizes, in the presence of NTF2-GST. We found that the fluxes of these control proteins are significantly impeded by the Nsp1FG membrane in a size-dependent manner, with fluxes of larger proteins being impeded more than those of smaller ones. At the same time, the flux of NTF2-GST in these paired mixtures exhibited no discernable reduction (Supplementary Table 3), showing that membranes functionalized with Nsp1FG behave as a selective filter with a strong preference for a transport factor (NTF2).

We next asked whether the Nsp1FG-functionalized membranes had selectivity for other transport factors. Whereas NTF2 is the major transport factor for the Ras-related nuclear protein Ran, most transport factors belong to a different structurally related family that are collectively termed karyopherins (Kaps). Karyopherins that mediate import bind to specific 'nuclear localization signals' in their cargos and, by virtue of their ability to bind FG-nucleoporins, facilitate passage of these cargos through the NPC into the nucleus. We studied two budding yeast karyopherins in our device, Kap95 and Kap121 (also known as Pse1). Both karyopherins have been used as standard reporters for nuclear import, and their specific interactions with

Nsp1 are well documented<sup>18,19</sup>. Similar to NTF2-GST, we found that these karyopherins pass essentially unimpeded through the Nsp1FG membranes compared with the control membranes. At the same time, BSA is again significantly impeded to the same degree as in the NTF2-GST experiments (Fig. 2b and Supplementary Table 3). Hence, the selective gating of the Nsp1FG-functionalized membranes seems to be general for FG-repeat-binding transport factors.

We also tested whether transport factors can carry their cargo molecules efficiently through the FG-nucleoporin-coated membranes, as they do through the NPC *in vivo*. To make a fluorescent cargo, we attached the Kap95-binding nuclear localization signals (the 'Ibb' (importin- $\beta$  binding domain of Kap60) sequence) from Kap60 onto green fluorescent protein (Ibb-GFP), with GFP acting as a control. We used a twofold excess of unlabelled Kap95 over the cargo; as expected, all Ibb-GFP was bound to Kap95 (here designated Kap95-Ibb-GFP) under these conditions, whereas GFP did not bind to Kap95 (Supplementary Fig. 9)<sup>20,21</sup>. Like Kap95 alone, the Kap95-Ibb-GFP complex was efficiently transported across the membranes with a flux ratio of  $\sim$ 1.0. In contrast, GFP without the Ibb domain and Ibb-GFP (in the presence of NTF2-GST but without Kap95, Supplementary



**Figure 2 | Selective trafficking of transport factors and cargo through the nanopores.** A flux ratio was obtained for each protein by calculating the ratio of its flux through a functionalized membrane (in this case Nsp1FG) relative to its flux through the control small PEG membrane. A value of 1 indicates no reduction in flux through the functionalized membrane as compared with the control membrane, whereas a value of 0 indicates no flux through the functionalized membrane. **a**, Flux ratios of NTF2–GST versus control proteins. Flux ratios plotted against Stokes radius for variously sized proteins in the presence of NTF2–GST ( $R_s = 3.6$  nm), namely: RNase A (1.75 nm), GFP (2.8 nm), BSA (3.5 nm), transferrin (3.6 nm), immunoglobulin G (IgG; 5.1 nm). The flux ratio of all proteins except NTF2–GST was reduced in a size-dependent manner. **b**, Flux ratios of two karyopherins versus control protein (BSA). Flux ratios for a mixture of BSA and Kap95 and a mixture of BSA and Kap121 showed BSA having a markedly lower flux ratio than either karyopherin (despite the smaller size of BSA) in a manner similar to a mixture of BSA and NTF2–GST. **c**, Karyopherin-mediated transport of cargo. Flux ratios for either GFP or Ibb–GFP in the presence of Kap95 (not fluorescently labelled) and IgG; Ibb–GFP forms a complex with Kap95, whereas GFP does not (see Supplementary Information). The flux ratio of Ibb–GFP in a complex with Kap95 was higher than that of either GFP alone (even though GFP has a much smaller Stokes radius than the Ibb–GFP–Kap95 complex) or IgG alone. Hence, the device mimics Kap-mediated cargo transport. Standard errors are shown.

Fig. 9c) had a markedly lower flux ratio of  $\sim 0.55$  and  $0.44$ , respectively, even though they have significantly smaller Stokes radii than the transport-factor–cargo complex (Fig. 2c). Similar results were obtained with NTF2–GST and its cargo, RanGDP (Supplementary Table 3). We have thus reconstituted the transport-factor-mediated import step of nuclear transport in our device.

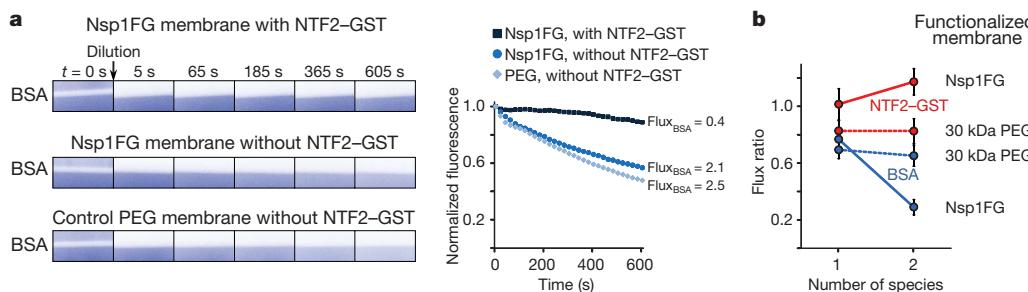
Reversible binding to FG-nucleoporins is an essential feature of nuclear transport, allowing transport factors to transiently interact with multiple docking sites as they traverse the NPC<sup>5</sup>. This feature is recapitulated in the present functionalized membrane, because we observed a specific and reversible accumulation of transport factors at the FG layers (for example, Fig. 1d and Supplementary Fig. 10a).

We next assayed what happens to the flux of the control proteins across Nsp1FG-coated membranes in the absence of transport factors (Fig. 3), measuring the flux of fluorescently labelled BSA in either the absence or the presence of NTF2–GST. In the absence of NTF2–GST the flux of BSA was only modestly impeded by the Nsp1FG-coated

membrane when compared to the small-PEG-coated membrane. Notably, the addition of NTF2–GST caused a substantial reduction in the flux of BSA (Fig. 3a); similar results were obtained for transferrin (Supplementary Tables 2 and 3). To test whether it was the binding of NTF2–GST to the Nsp1FG layer that caused this marked reduction in BSA flux, we repeated the experiment using membranes functionalized with PEG thiol of  $\sim 30$  kDa (30 kDa PEG), which binds neither BSA nor NTF2–GST but is similar to Nsp1FG in both its size and its behaviour as an unstructured polymer<sup>9,22,23</sup>. Both NTF2–GST and BSA showed only moderately reduced fluxes through the 30 kDa PEG membrane as compared to the control; moreover, the presence or absence of NTF2–GST had no effect on the flux of BSA (Fig. 3b). We conclude that the strong flux reduction of control proteins occurs only in the presence of FG-nucleoporin-binding transport factors—that is, addition of transport factors markedly changes the transport properties of the functionalized membrane. Possible explanations for this effect include direct competition between transport factors and other macromolecules for space and binding sites inside the NPC<sup>24</sup> and changes in the configurations of the FG-nucleoporin barrier on transport factor binding (see for example ref. 8). These effects are not necessarily mutually exclusive.

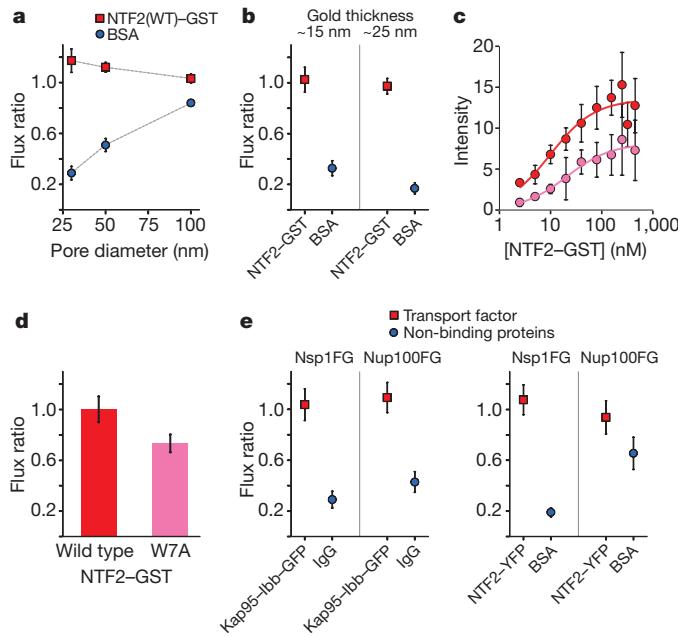
We predicted that an increase in pore size should reduce the concentration of FG-nucleoporins in the pore, thus reducing the selectivity of the functionalized membrane. Indeed, we found that the larger the pore size, the weaker the selectivity of the membrane (Fig. 4a). We also predicted that an increase in the thickness of the gold layer should allow more FG-nucleoporins to bind to the internal gold surface at the entrance to the polycarbonate pores, while leaving the flat gold surface unchanged (Fig. 1a). We observed that a  $\sim 2$ -fold increase in gold layer thickness yielded a  $\sim 2$ -fold increase in selectivity (Fig. 4b), without changing the effective pore size (Supplementary Table 6). This result also shows how future carefully tuned alterations in the design of our device could significantly enhance its selectivity.

Because binding of transport factors to FG-nucleoporins is a central feature of transport selectivity, we assessed the effect of mutations that alter the binding strength of NTF2 to Nsp1. We used mutant NTF2 (NTF2(W7A)), which has a reduced interaction with FxFG nucleoporins due to an alteration in the main phenylalanine binding site of NTF2 (ref. 25). As with NTF2–GST, the NTF2(W7A)–GST mutant is a dimer (Supplementary Fig. 1), but has reduced binding capacity for Nsp1 (Fig. 4c; Supplementary Fig. 11a). In our device, NTF2–GST and NTF2(W7A)–GST were combined in equimolar ratios, and their transport through Nsp1FG and small PEG membranes was measured. As expected because of its reduced binding, the mutant showed a significantly reduced flux ratio compared to the wild type (Fig. 4d). Another way of changing binding is by changing the FG-repeat motif, because different kinds of motif have differing affinities for a given transport factor. For example, Nup100FG binds strongly to Kap95 but



**Figure 3 | Presence of transport factor enhances the selectivity of FG-nucleoporin-coated membranes.** **a**, Left, time course of confocal images collected as in Fig. 1, comparing fluxes of BSA through the small PEG membrane (bottom), the Nsp1FG membrane (middle), and the Nsp1FG membrane in the presence of NTF2–GST (top); right, corresponding fluorescence decrease curves and fluxes (in molecules pore<sup>−1</sup> s<sup>−1</sup>  $\mu\text{M}^{-1}$ ).

**b**, Strong selectivity of NTF2–GST transport over BSA is observed for the Nsp1FG functionalized membrane and is not observed for the ‘inert’ 30 kDa PEG functionalized membrane. Flux ratios are plotted for BSA and NTF2–GST alone (one species) or in combination with each other (two species), through either Nsp1FG or 30 kDa PEG membranes versus small PEG membrane. Standard errors are shown.



**Figure 4 | The effect of pore geometry and FG-nucleoporin binding strength on transport.** **a**, The selectivity of the Nsp1FG membranes decreases with increasing pore diameter. Flux ratio measurements of a mixture of BSA and NTF2-GST were performed as in Fig. 2a, for membranes having differing pore diameters. **b**, The selectivity of the Nsp1FG membrane was significantly improved when the thickness of the gold layer was increased. **c**, *In vitro*, NTF2(WT)-GST (red) had a lower apparent  $K_d$  value ( $10 \pm 3$  nM) than mutant NTF2(W7A)-GST (pink,  $K_d$  of  $21 \pm 4$  nM) for binding to Nsp1FG, and at the same time had about double the binding capacity compared to the mutant. **d**, The NTF2(W7A)-GST mutant had a reduced flux ratio compared to NTF2(WT)-GST (for flux through the Nsp1FG membrane,  $P = 0.009$ , single-tailed *t*-test), although no flux difference was seen on the control small-PEG-coated membranes ( $P = 0.35$ ). **e**, Assessment of the effect on transport of binding strength to FG-nucleoporins by changing the FG-repeat motif. Flux ratios through either Nsp1FG or Nup100FG membranes are shown for the Kap95-Ibb-GFP complex with IgG and for NTF2-YFP with BSA. Because NTF2 binds poorly to Nup100FG compared to its binding strength to Nsp1FG, Nup100FG-coated membranes did not discriminate well between BSA and NTF2-YFP. Standard errors are shown in parts **a**, **b**, **d** and **e**, and standard deviations are shown in **c**.

more weakly to NTF2 (refs 18, 26, 27; Supplementary Fig. 10). We therefore compared the flux ratios of Kap95 and NTF2-YFP through membranes functionalized with either Nsp1FG or Nup100FG. As expected, Kap95, which binds strongly to both Nsp1FG and Nup100FG, efficiently transits these membranes and impedes the transit of a control protein (Fig. 4e). Importantly, this result also demonstrates that other FG-nucleoporins behave in a similar manner to Nsp1 and can support selective transport in our device. In contrast, the flux of a control protein was reduced only slightly through the Nup100FG membranes when NTF2-YFP was present (Fig. 4e), again emphasizing that efficient binding is required for selective transport.

It has been shown that treatment by hexanediol abolishes the NPC barrier *in vivo*<sup>4,28</sup> and collapses FG-nucleoporins *in vitro*<sup>8</sup>. Consistent with these results, we observe that addition of the 5% hexanediol largely eliminates the selectivity for the transmission of NTF2-GST over BSA through an Nsp1FG membrane (Supplementary Fig. 8).

We have built a minimal assembly that encapsulates the essential functional features of NPC architecture as they are currently understood—namely, a simple pore of the correct dimensions coated with FG-nucleoporins. Notably, this assembly recapitulates key features of nucleocytoplasmonic transport, selectively discriminating against control proteins in favour of transport factors and transport-factor–cargo complexes. Our NPC mimic also provides insight

into the role of transport factors in the mechanism of selectivity. In a sense, transport factors may be considered transient components of the NPC that help to discriminate against the passage of nonspecific material across the nuclear envelope. A feature of our device is that it incorporates properties of nuclear transport that may prove useful for purification techniques in general, including an ability to extract macromolecules of interest most efficiently from crude mixtures. Such NPC-inspired molecular sorters could have many practical analytical and preparative applications.

## METHODS SUMMARY

**Protein expression and purification.** All proteins were expressed in *Escherichia coli*. His-tagged versions of Nsp1FG-Cys, Nup100FG-Cys and human NTF2-yellow fluorescent protein (YFP) were purified from clarified lysate using Ni- or Co-affinity chromatography followed by gel filtration chromatography. Human wild-type NTF2-GST and NTF2(W7A)-GST were purified using glutathione sepharose resin and labelled with Cy dyes according to the manufacturer's instructions (GE Healthcare). Kap95 and Kap121 were obtained as described previously<sup>29</sup> and labelled with an Alexa 488 protein-labelling kit according to the manufacturer's instructions (Invitrogen). His-tagged versions of Ibb-GFP and GFP were purified using Co-affinity resin.

**Construction of the functionalized membrane.** Gold-sputtered polycarbonate membranes (GE Osmomics) were cleaned with 25% HNO<sub>3</sub> and rinsed in water. Chemisorption of 6.5 μM Nsp1FG-Cys, 5.1 μM Nup100FG-Cys, 2 mM PEG-thiol (30 kDa) or 2 mM PEG-thiol (356 Da) was performed at room temperature (22 °C) for 1 h. Protein-coated membranes were subsequently incubated with 2 mM PEG-thiol (356 Da) in 8 M urea for 1 h and all membranes were washed in TBT buffer (20 mM HEPES, pH 7.4, 110 mM KOAc, 2 mM MgCl<sub>2</sub>, 10 μM ZnCl<sub>2</sub>, 10 μM CaCl<sub>2</sub>, 0.1% Tween-20). Amino acid analysis on the protein-coated membranes confirmed specific chemisorptions of 0.05 and 0.04 molecules of Nsp1FG and Nup100FG per nm<sup>2</sup>, respectively.

**Fluorescence imaging.** Fluorescence images were recorded on a Leica TCS SP spectral confocal microscope with a PL × 100/1.4 oil-immersion objective lens in XZT mode. Mean fluorescence values were obtained using ImageJ, normalized, and fit using a simple exponential decay (SigmaPlot) yielding the flux (molecules pore<sup>-1</sup> s<sup>-1</sup> μM<sup>-1</sup>). Full Methods are provided in the Supplementary Information.

Received 12 September; accepted 29 October 2008.

Published online 21 December 2008.

- Alber, F. et al. The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701 (2007).
- Macara, I. G. Transport into and out of the nucleus. *Microbiol. Mol. Biol. Rev.* 65, 570–594 (2001).
- Peters, R. Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality. *Traffic* 6, 421–427 (2005).
- Ribbeck, K. & Gorlich, D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* 21, 2664–2671 (2002).
- Rout, M. P., Aitchison, J. D., Magnasco, M. O. & Chait, B. T. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* 13, 622–628 (2003).
- Rout, M. P. et al. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148, 635–651 (2000).
- Frey, S. & Gorlich, D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* 130, 512–523 (2007).
- Lim, R. Y. H. et al. Nanomechanical basis of selective gating by the nuclear pore complex. *Science* 318, 640–643 (2007).
- Lim, R. Y. H. et al. Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmonic transport. *Proc. Natl. Acad. Sci. USA* 103, 9512–9517 (2006).
- Krishnan, V. et al. Intramolecular cohesion of coils mediated by phenylalanine-glycine motifs in the natively unfolded domain of a nucleoporin. *PLOS Comput. Biol.* 4, e1000145 (2008).
- Lakshmi, B. B. & Martin, C. R. Enantioseparation using apoenzymes immobilized in a porous polymeric membrane. *Nature* 388, 758–760 (1997).
- Lee, S. B. et al. Antibody-based bio-nanotube membranes for enantiomeric drug separations. *Science* 296, 2198–2200 (2002).
- Jirage, K. B., Hulteen, J. C. & Martin, C. R. Effect of thiol chemisorption on the transport properties of gold nanotube membranes. *Anal. Chem.* 71, 4913–4918 (1999).
- Iqbal, S. M., Akin, D. & Bashir, R. Solid-state nanopore channels with DNA selectivity. *Nature Nanotechnol.* 2, 243–248 (2007).
- Caspi, Y., Zbaida, D., Cohen, H. & Elbaum, M. Synthetic mimic of selective transport through the nuclear pore complex. *Nano Lett.* 8, 3728–3734 (2008).
- Strawn, L. A., Shen, T. X., Shulga, N., Goldfarb, D. S. & Wente, S. R. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nature Cell Biol.* 6, 197–206 (2004).

17. Peters, R. Optical single transporter recording: transport kinetics in microarrays of membrane patches. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 47–67 (2003).
18. Bayliss, R., Littlewood, T., Strawn, L. A., Wente, S. R. & Stewart, M. GLFG and FxFG nucleoporins bind to overlapping sites on importin- $\beta$ . *J. Biol. Chem.* **277**, 50597–50606 (2002).
19. Seedorf, M., Damelin, M., Kahana, J., Taura, T. & Silver, P. A. Interactions between a nuclear transporter and a subset of nuclear pore complex proteins depend on Ran GTPase. *Mol. Cell. Biol.* **19**, 1547–1557 (1999).
20. Koerner, C., Guan, T., Gerace, L. & Cingolani, G. Synergy of silent and hot spot mutations in importin beta reveals a dynamic mechanism for recognition of a nuclear localization signal. *J. Biol. Chem.* **278**, 16216–16221 (2003).
21. Gilchrist, D., Mykytka, B. & Rexach, M. Accelerating the rate of disassembly of karyopherin-cargo complexes. *J. Biol. Chem.* **277**, 18161–18172 (2002).
22. Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L. & Rexach, M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl Acad. Sci. USA* **100**, 2450–2455 (2003).
23. Dutta, A. K. & Belfort, G. Adsorbed gels versus brushes: viscoelastic differences. *Langmuir* **23**, 3088–3094 (2007).
24. Zilman, A., Di Talia, S., Chait, B. T., Rout, M. P. & Magnasco, M. O. Efficiency, selectivity, and robustness of nucleocytoplasmic transport. *PLOS Comput. Biol.* **3**, 1281–1290 (2007).
25. Morrison, J., Yang, J. C., Stewart, M. & Neuhaus, D. Solution NMR study of the interaction between NTF2 and nucleoporin FxFG repeats. *J. Mol. Biol.* **333**, 587–603 (2003).
26. Clarkson, W. D. et al. Nuclear protein import is decreased by engineered mutants of nuclear transport factor 2 (NTF2) that do not bind GDP-Ran. *J. Mol. Biol.* **272**, 716–730 (1997).
27. Strawn, L. A., Shen, T. X. & Wente, S. R. The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex. *J. Biol. Chem.* **276**, 6445–6452 (2001).
28. Shulga, N. & Goldfarb, D. S. Binding dynamics of structural nucleoporins govern nuclear pore complex permeability and may mediate channel gating. *Mol. Cell. Biol.* **23**, 534–542 (2003).
29. Leslie, D. M. et al. Characterization of karyopherin cargoes reveals unique mechanisms of kap121p-mediated nuclear import. *Mol. Cell. Biol.* **24**, 8487–8503 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank E. Coutavas, S. Darst, G. Blobel and C. Martin for suggestions and comments, G. Blobel for use of his confocal microscope, D. Phillips for use of his sputtering device, P. Nahirney and A. Labissiere for electron microscopy work, J. M. Crawford for amino acid analysis, D. Gadsby and A. Gulyas Kovacs for providing *Xenopus* oocytes, J. Aitchison for Kap95-GST and Kap121-GST plasmids, K. Zerf and M. Kahms for providing RanGDP, K. Zerf for NTF2-YFP cloning assistance, R. Mironski for help in preparing measuring chambers, and other members of the Peters, Rout and Chait laboratories for their assistance. We gratefully acknowledge support from the NIH and DoE. J.T.-N. is a HHMI pre-doctoral fellow.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to B.T.C. ([chait@rockefeller.edu](mailto:chait@rockefeller.edu)).

## LETTERS

# Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs

Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project\*

The transcriptomes of eukaryotic cells are incredibly complex. Individual non-coding RNAs dwarf the number of protein-coding genes, and include classes that are well understood as well as classes for which the nature, extent and functional roles are obscure<sup>1</sup>. Deep sequencing of small RNAs (<200 nucleotides) from human HeLa and HepG2 cells revealed a remarkable breadth of species. These arose both from within annotated genes and from unannotated intergenic regions. Overall, small RNAs tended to align with CAGE (cap-analysis of gene expression) tags<sup>2</sup>, which mark the 5' ends of capped, long RNA transcripts. Many small RNAs, including the previously described promoter-associated small RNAs<sup>3</sup>, appeared to possess cap structures. Members of an extensive class of both small RNAs and CAGE tags were distributed across internal exons of annotated protein coding and non-coding genes, sometimes crossing exon-exon junctions. Here we show that processing of mature mRNAs through an as yet unknown mechanism may generate complex populations of both long and short RNAs whose apparently capped 5' ends coincide. Supplying synthetic promoter-associated small RNAs corresponding to the c-MYC transcriptional start site reduced MYC messenger RNA abundance. The studies presented here expand the catalogue of cellular small RNAs and demonstrate a biological impact for at least one class of non-canonical small RNAs.

The repertoire of RNAs found in eukaryotic cells is unexpectedly complex, with virtually the entire non-repeat portions of many genomes being transcribed<sup>1</sup>. Genic regions are often populated by interleaved transcription units, which give rise to both protein-coding RNAs and long and short non-coding RNAs<sup>1</sup>. Promoter-associated short RNAs (PASRs) and termini-associated short RNAs (TASRs) are recent additions to the pantheon of short RNAs<sup>4</sup>. Although their functions are unknown, several of their characteristics support biological significance. For example, PASRs and TASRs cluster at 5' and 3' termini of annotated genes<sup>4</sup>. Overall, the presence of PASRs correlates with the expression of a given locus, but not all expressed loci generate these species. Moreover, the production of PASRs and TASRs from particular loci is a conserved feature of the human and mouse genomes<sup>4</sup>. As part of our ongoing effort to understand the full repertoire of small RNAs, their mechanisms of biogenesis and their biological impacts, we analysed the small RNAs (<200 nucleotides (nt)) of HepG2 and HeLa cell lines using next-generation sequencing<sup>5</sup>.

Nearly 80 million short sequence reads (30–35 bases) were generated, representing RNAs <200 nt in both cell lines (Supplementary Fig. 1). Our sequencing protocols favoured RNAs with 5' mono-, di- and tri-phosphate groups and capped RNAs. Nearly 30 million of these could be matched perfectly to the hg18 release of the human genome, with 9.5 million reads mapping to unique sites (Supplementary Fig. 1).

Sequences derived from mitochondria, chromosome Y, repeats, annotated small RNAs, predicted RNA genes<sup>6</sup>, and known and

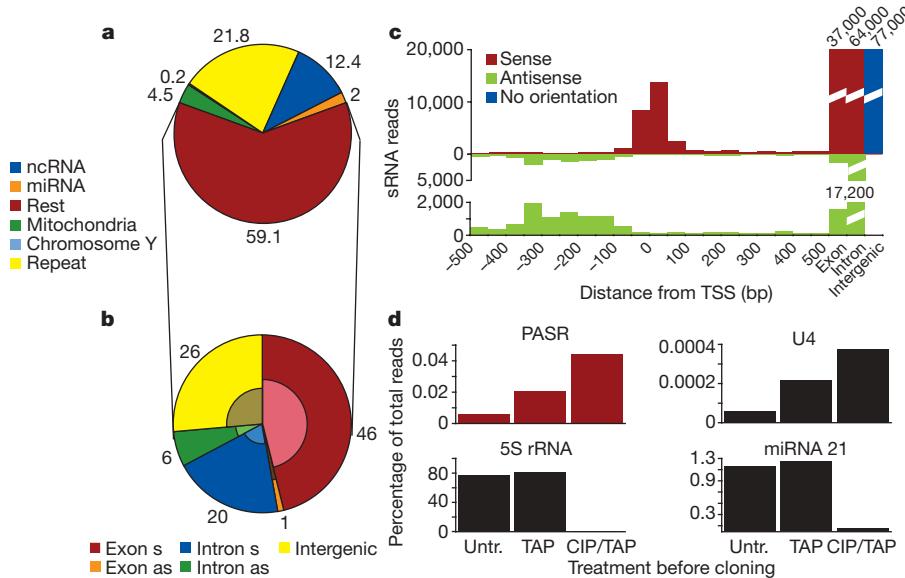
predicted small nucleolar RNA (sno)RNAs<sup>7</sup> were excluded from further analysis (Fig. 1a and Supplementary Fig. 1). This resulted in 232,805 sequences representing new small RNAs. Independent libraries from the same cell line only modestly overlap, indicating that our studies have not saturated small RNA content. Sequences were collapsed based on their mapping positions, assigned as the 5' nucleotide of each read. This resulted in 102,159 distinct 5' ends ('rest', Fig. 1a). A large fraction of sequences are derived from unannotated intergenic regions (Fig. 1b). Sequences also matched the sense and antisense strands of exonic and intronic regions of annotated genes. Notably, nearly half of all reads could be assigned to the sense strand of annotated exons, with a strong representation of first exons.

We had previously noted a class of small RNAs, namely PASRs, which associated with transcriptional start sites (TSSs) and mapped both to promoter regions and to annotated first exons. We therefore plotted the distribution of the unannotated RNA category (rest, Fig. 1a, b) with respect to known TSSs. A clear pattern emerged with an enrichment of small RNAs on both strands directly adjacent to the TSS (Fig. 1c). The sense and antisense strands surrounding TSSs showed mirror-image profiles, with small RNAs on the sense strand accumulating most strongly downstream of the annotated TSS and small RNAs on the antisense strand accumulating mainly upstream of the annotated TSS. This is similar to what had been previously observed on high-resolution genomic tiling arrays<sup>4</sup>. A gap of ~50 nt on the antisense strand separated the precise TSS and PASRs, an observation we have yet to understand (Fig. 1c).

This PASR class (previously defined as small RNAs mapping within 500 nt of an annotated TSS) is found in both genic and intergenic regions of the genome and comprises 16.2% or 17.7% of filtered sequence tags (un-collapsed or collapsed, respectively). PASRs contribute to the small RNAs placed within both strands of exons and introns and small RNAs annotated as intergenic (Fig. 1b, shaded inner circles in each pie chart). On the basis of these definitions, PASRs form the most abundant individual class of defined small RNAs within the non-annotated fraction of our sequences.

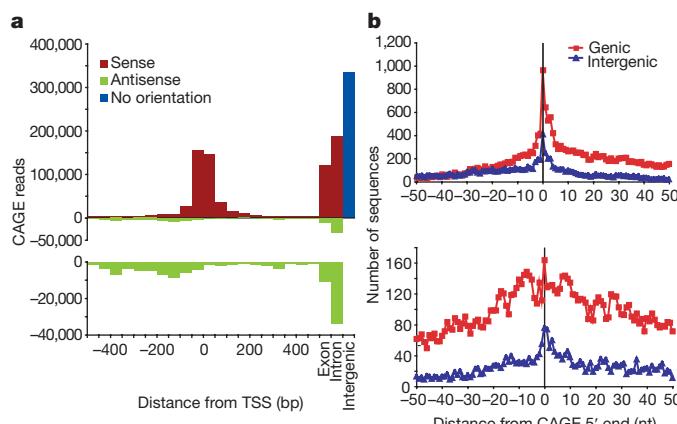
Because PASRs strongly associate with TSSs, we posited that transcription initiation *per se* might generate PASR 5' ends. Thus, PASRs might contain cap structures. To probe this possibility, we prepared small RNA libraries using methods that require the presence of a 5' phosphate and 3' hydroxyl group<sup>8</sup>. Capped RNAs should be refractory to capture by this cloning protocol, but could be made susceptible to cloning by removing cap structures. We therefore prepared three different small RNA libraries from HepG2 cells. One was from untreated RNA. The second was from RNA treated with tobacco acid pyrophosphatase (TAP) to leave clonable monophosphorylated ends on RNAs with caps, or with di- or tri-phosphate termini. The third library was from RNA treated with calf intestinal alkaline phosphatase (CIP) before TAP treatment. Pre-treating with CIP removes phosphates to leave unclonable 5' OH termini on all uncapped

\*Lists of participants and their affiliations appear at the end of the paper.



RNAs. Sequence tags corresponding to the 5' end of the capped U4 small nuclear (sn)RNA are enriched in libraries by TAP treatment and further enriched by CIP addition before TAP treatment (Fig. 1d). MicroRNA 21, which has a 5' monophosphate terminus, is lost from the library on CIP treatment, as is 5S ribosomal RNA, which has a 5' triphosphate terminus (Fig. 1d). Small RNAs defined as PASRs follow the pattern established by U4, consistent with them bearing some type of cap structure. The observation that PASRs are revealed to the cloning protocol by TAP alone indicates that they also contain 3' OH termini. Considered together, these data indicate that PASRs are likely to arise either as independent capped transcripts emanating from annotated TSSs on both genomic strands or as processing products from longer capped RNAs. A candidate for the latter are PALRs (promoter-associated long RNAs), which often extend through the first exon and into the first intron<sup>4</sup>.

CAGE tagging protocols take advantage of the 5' cap structure to capture sequence reads from the 5' ends of long RNAs<sup>2</sup>. Substantial databases of such tags have been produced from long polyadenylated RNAs from more than 20 human tissues<sup>9,10</sup>. Plotting CAGE tags with respect to annotated TSSs revealed patterns similar to those observed for PASR class small RNAs (Fig. 2a). In both genic and intergenic regions, we also observed a strong tendency for a precise identity between CAGE and PASR 5' ends (Fig. 2b).



**Figure 2 | Correlation of sRNAs and CAGE tags.** **a**, Distribution of CAGE tags over annotated TSSs. Orientation is with respect to the long transcript. Antisense sRNAs are plotted with a different y-axis beneath. (Uncollapsed data.) **b**, Distribution of PASR (top) and non-PASR sRNAs (bottom) around CAGE tag 5' ends. The distance to closest short RNA 5' end was plotted for each CAGE tag. (Collapsed data.)

### Figure 1 | Genomic distribution of small RNAs.

**a**, Annotation of sRNAs from sequencing. 'Rest' represents unannotated sRNAs filtered for mitochondria, chromosome Y, repeats and known sRNAs. miRNA, microRNA; ncRNA, non-coding RNA. (Collapsed data are shown.)

**b**, Mapping of unannotated sequences to annotated genomic landmarks. *as*, antisense to corresponding transcript; *s*, sense. (Collapsed data.) Shaded inner portions represent the fraction that are PASRs. **c**, Distribution of sRNAs over TSSs. Orientation is with respect to the long transcript. Antisense sRNAs are plotted with a different y-axis beneath. (Uncollapsed data.)

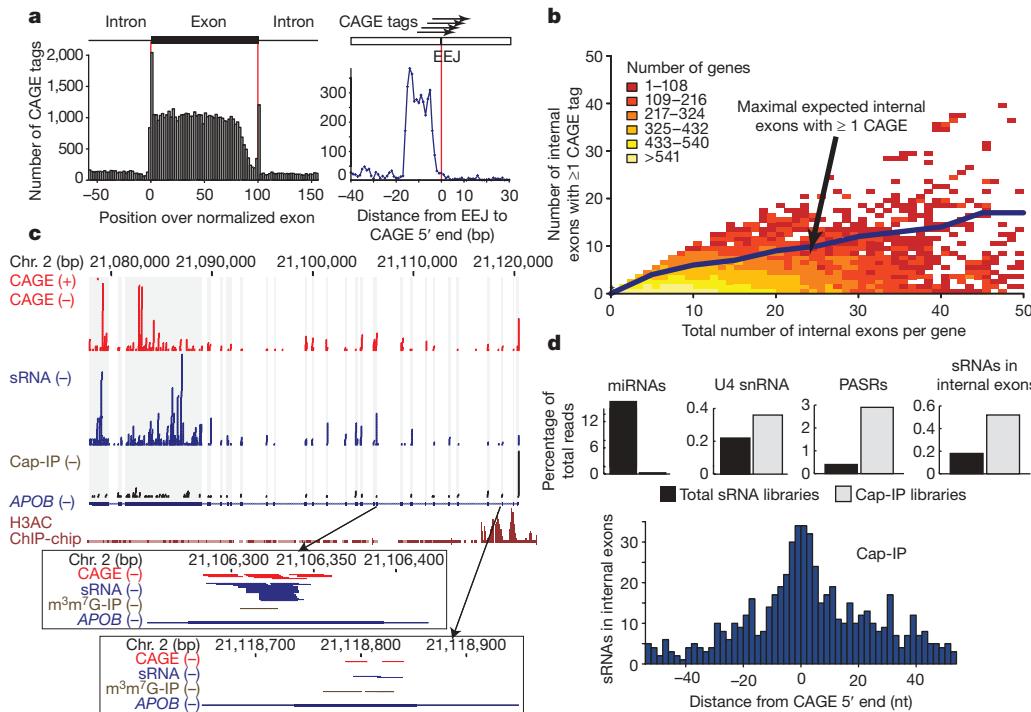
**d**, Characterization of PASR 5' ends. *Untr.*, untreated. Sequences corresponding to the 5' end of U4, 5S rRNA and mir-21 were extracted as controls. (Uncollapsed data.)

We also noted a substantial population of small RNAs mapping more than 500 nt away from annotated TSSs (Fig. 1c), which contributed to intronic, exonic and intergenic classes (Fig. 1b, outer portions of each pie). Similarly we noted a large population of CAGE tags located >500 base pairs (bp) from the TSS (59.0%) that could be assigned to exons, introns and intergenic regions (11.36, 18.8 and 28.9% of total uncollapsed CAGE tags, respectively; Supplementary Table 1).

Certainly, a fraction of these could arise as products from unannotated TSSs, giving rise to both CAGE tags and PASRs. However, the correlation between the 5' ends of non-PASR small RNAs and CAGE tags was less precise ( $\pm 10$  nt) than was noted at annotated TSSs (Fig. 2b). This indicated that non-PASR classes of small RNAs might arise by mechanisms that differ from canonical PASRs.

Both small RNAs and CAGE tags accumulate more strongly in internal exons than in introns or in intergenic space, if these regions are normalized by their cumulative length (Fig. 3a and not shown). By examining the distribution of both CAGE tags and small RNAs across annotated internal exons, we noted a strongly decreasing number of CAGE sequences beginning about 20 nt from the 3' end of the average exon/intron boundary (splice donor site; Fig. 3a). If CAGE tags crossed splice junctions, they would not be co-linear with the genome at these sites and would, therefore, not have been mapped in the initial analyses<sup>10</sup>, possibly giving rise to the observed pattern. We therefore extracted previously published CAGE tags, which had failed to map to the genome, and probed these against sequences of known exon-exon junctions. We uncovered a substantial population of CAGE tags that crossed splice junctions, and that therefore must have arisen from at least partially processed mRNAs (Fig. 3a).

CAGE tags are well established as markers of capped 5' ends. Certainly, internal exons might contain unknown sites of transcriptional initiation that could give rise to both CAGE tags and small RNAs, which would then be defined as PASRs. However, we observe numerous tags that both initiate less than 20 bases from exon boundaries and cross exon-exon junctions. Very short exons splice inefficiently<sup>11</sup>, and naturally occurring 5' exons less than 20 bases in length are rare (not shown). Thus, the CAGE tags that we observe probably represent cleaved products of mature mRNAs that somehow acquire a 5' modification analogous to a cap structure that renders them sensitive to the CAGE tagging method. Such a reaction would represent a previously unrecognized RNA processing pathway and a previously unknown fate for spliced mRNAs. Although the CAGE tags used for comparison in this study were derived from polyadenylated RNAs, we cannot determine whether small RNAs originated from poly(A)<sup>+</sup> or poly(A)<sup>-</sup> transcripts.



**Figure 3 | Correlation between CAGE tags, sRNAs and internal exons of annotated transcripts.** **a**, Left: distribution of mapped CAGE tag 5' ends across internal exons. Exon length was normalized to 100 segments. Right: distribution of CAGE tags not mapping to the genome but mapping to exon–exon junctions (EEJ) of internal exons. **b**, Prevalence of internal CAGE tags. Black line represents the maximum expected exons in random samplings (see Methods). Colour corresponds to number of transcripts represented by each data point. **c**, CAGE tag and sRNA coverage of the *APOB* gene. **d**, sRNAs from cap-immunoprecipitation (IP) are shown separately. Histone H3 acetylation (H3AC) pattern<sup>13</sup> is shown below. Two internal exons are magnified. **d**, Characterization of libraries from anti-cap-immunoprecipitated RNA. Top panel: representation of sRNAs in total and IP libraries (uncollapsed data). For all but the U4 fraction, uniquely mapping sequences were considered. Bottom panel: distance to closest sRNA 5' end from CAGE 5' end in internal exons (collapsed data).

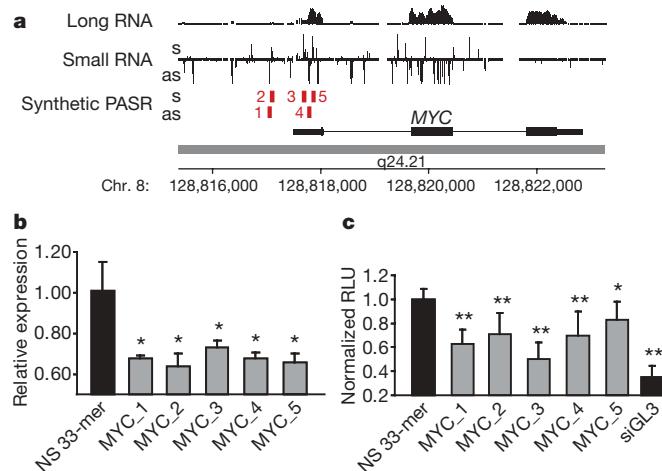
Generation of CAGE tags from internal exons is not confined to a small number of genes. In fact, 49% of all human genes generate a CAGE tag mapping to an internal exon. For 2% of them, one or more CAGE tags are found in all of the internal exons (Fig. 3b). This exceeds the number expected by chance ( $P$  value  $< 0.001$ , Fig. 3b). Prevalent and systematic generation of both small RNAs and CAGE tags from internal exons is illustrated by *APOB* (Fig. 3c), a gene encoding the apolipoprotein B protein that regulates cholesterol metabolism<sup>12</sup>. In this instance, mapping of acetylated histone H3 (ref. 13) is consistent with a single prevalent TSS, which correlates with the presence of both CAGE tags and PASRs (Fig. 3c). However, both CAGE tags and small RNAs are even more abundant in internal *APOB* exons, and these often coincide at specific sites (Fig. 3c, insets). This provokes a model in which mature transcripts from the *APOB* gene are processed post-transcriptionally and in which processing products become modified by some type of cap structure. This possibility gains support from sequencing of small RNAs recovered by immunoprecipitation with a methylguanosine cap antibody<sup>14</sup>. This not only recovers PASRs but also enriches small RNAs that map less than 10 bp away from a CAGE tag in internal exons ( $P$  value  $< 0.001$ , see also Fig. 1d). In libraries prepared from unfractionated small RNAs, 10.0% of filtered sequences mapping within internal exons lie within 10 bp of a CAGE tag. This number increases to 27.0% in libraries prepared from RNAs immunoprecipitated with the anti-m<sup>7</sup>G antibody.

As illustrated by *APOB*, genic CAGE tags and small RNAs are approximately ten times more likely to map within exonic than intronic regions (Fig. 3a, b). As with CAGE tags crossing exon–exon junctions, this result is consistent with a model in which CAGE tags can be derived from products of processed mRNAs.

Considering the prevalence of the PASR class, we sought to probe its potential biological function. As with many genes, PASRs are found at the annotated TSSs of the *MYC* oncogene<sup>4</sup>. We synthesized a collection

of 30–35-nt, single-stranded RNAs that share their 5' ends with three PASRs from the sense genomic strand and two from antisense strand upstream of the annotated TSS (Fig. 4a). These were transfected individually into HeLa cells, and their effect on the abundance of *MYC* mRNA was measured (Fig. 4b). In each case, transfection of the PASR mimetic reduced the expression of *c-MYC* mRNA. The consequences of these effects were measured by co-transfection of a *MYC*-responsive luciferase reporter construct, which showed reduced activity in the presence of each PASR (Fig. 4c). Similar results were obtained for five PASR mimetics corresponding to the connective tissue growth factor (*CTGF*) gene (Supplementary Fig. 2). The presence of PASRs is associated with marks of active transcription, including association with RNA polymerase II, histone H3 and H4 acetylation, and H3K4 tri-methylation, as well as an increased susceptibility to DNase treatment (Supplementary Fig. 3). Our data indicate a causal connection between PASRs and active *MYC* expression, although we have not yet investigated the impact of delivering ectopic PASRs on the active marks with which the presence of endogenous species is correlated.

Profiling of small RNAs, defined as those less than 200 nt in length, has revealed a substantial complexity in the output of both genic and intergenic regions of the genome. These studies have raised two possibilities for the origin of PASRs. First, they may be produced as capped, independent transcription products from promoters that also generate long RNAs. Second, they may be generated as post-transcriptional processing products of longer RNAs that initiate at annotated TSSs. A notable outcome of these studies is the finding that both CAGE tags and small RNAs decorate not only intergenic spaces but also internal exons of protein coding and non-coding transcripts. The existence of a large class of CAGE tags that are both adjacent to and cross splice junctions provides a *prima facie* case for the conclusion that long RNAs are metabolized into short RNAs that bear cap-like structures at their 5' ends (Fig. 5). The long RNAs, which



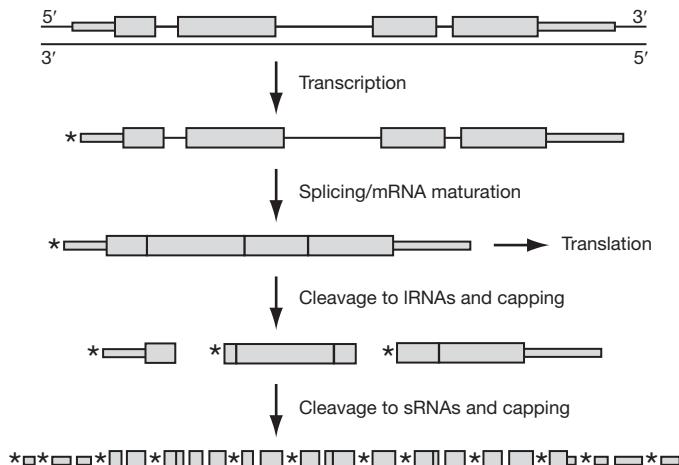
**Figure 4 | Regulation of gene expression by PASRs.** **a**, Expression profile of the *MYC* locus. The long and short RNA profile of HeLa cells on Affymetrix tiling arrays. Red rectangles indicate the designed synthetic PASRs (MYC\_1–5 are denoted by numbers and sequence information is provided in Supplementary Table 2) corresponding to peaks in the sRNA array profile. **b**, *MYC* mRNA expression levels in HeLa cells as measured by quantitative PCR with reverse transcription ( $n = 3$ ,  $P$  values  $< 0.01$ ). **c**, Effects of PASR transfections on a *MYC*-responsive luciferase transcriptional reporter in HeLa cells was measured as relative light units (RLU) ( $n = 2$ , \* $P < 0.01$ , \*\* $P < 0.001$ ). For reference, a control 33-mer and an siRNA directed against luciferase (siGL3) are shown.

ultimately give rise to short RNAs, could be primary transcription products or processing products themselves. Moreover, our studies indicate that CAGE tags are capturing not only TSSs but also the 5' ends of post-transcriptionally processed RNAs.

A key question remains as to whether the group of small RNAs that arise from internal exons represents transition products from mature mRNAs into recyclable ribonucleotides. Several lines of evidence argue against these representing simple degradation intermediates. First, there is a strong correlation between the precise 5' ends of CAGE tags, derived from long RNAs, and small RNAs identified in our study. These maps were produced from various RNA and tissue sources, and by different groups. The results from independent samples are consistent and indicative of discrete processing sites. Second, based on chemical modification in the CAGE procedure and affinity purification for small RNA libraries, both types of tags significantly enrich under conditions that favour capped RNAs. Third, CAGE tags and small RNA species arise only from a discrete, although substantial, subset of genes, and the abundance of the non-PASR class does not correlate simply with the expression level of their generative loci (see Supplementary Methods).

Several studies have indicated that RNA interference directed to promoter regions and apparently non-transcribed portions of genes can have a regulatory impact. In some cases that impact is silencing<sup>15,16</sup>, whereas in others activation was surprisingly observed<sup>17</sup>. Our analysis of PASRs indicates that providing their synthetic mimetics *in trans* can have a consistent, although modest, impact on gene expression. Although in the two cases tested, *MYC* and *CTGF*, increasing PASR levels decreased expression, it remains possible that the outcome of manipulating PASRs will be gene-specific, consistent with accumulating evidence that destroying promoter-associated RNA (PASR) species can have both positive and negative impacts<sup>17</sup>.

The functions of small RNAs corresponding to intergenic and exonic regions remain obscure. Such species could have regulatory roles, *per se*, or they could participate more globally in a bookkeeping or quality control mechanism by which the cell records its transcriptional output and splicing patterns. This has been previously hypothesized as a role for non-protein-coding RNAs<sup>18,19</sup>. What is clear is that the transcriptional product of cells is captured in small, stable RNA populations to a



**Figure 5 | A proposed model for the metabolism of genic transcripts into a diversity of long and short RNAs.** Transcription of a genic region results in a precursor long RNA containing a 5' cap structure, as shown by asterisks. After processing into spliced RNAs, protein-coding RNAs are destined either to be translated or to be further processed. This further processing entails cleavage followed in some cases by addition of a 5' modification, possibly a cap structure. Additional cleavage of these intermediate products can generate a class of short RNAs, some also bearing a cap structure. IRNAs, long RNAs.

degree that was unanticipated, and that at least a subset of these can serve as components of regulatory circuits.

## METHODS SUMMARY

A detailed description of the methods is provided in Supplementary Information. Short RNA was extracted from HeLa and HepG2 cells and subjected to CIP and TAP treatments as indicated. Adaptor sequences were added by ligation at the 5' end and by C-tailing and reverse transcription at the 3' end. Libraries were PCR-amplified and sequenced on an Illumina GA2. RNAs bearing 5' caps were enriched by immunoprecipitation using an m<sup>3</sup>/m<sup>7</sup>G antibody.

Analysis was performed on the hg18 assembly of the human genome using UCSC annotations. CAGE tag sequences were from the RIKEN institute website<sup>9</sup>. For short (s)RNA density, the distance to the closest annotated TSS or CAGE tag was calculated. Internal exons that do not overlap initial or terminal exons from the UCSC annotation were retrieved and distances from the acceptor sites to the mapped CAGE tags were normalized to the length of the corresponding exon. CAGE tags not mapping to the genome were mapped to exon-exon junctions containing the last and the first 50 bp of consecutive exons. To determine the prevalence of genes with internal CAGE tags, the number of internal exons containing at least one CAGE was plotted against the total number of exons in each gene. Transfection experiments were performed similar to those used for siRNAs<sup>20</sup> using single-stranded RNA oligonucleotides. Messenger RNA levels were determined 48 h later using quantitative PCR with reverse transcription. The *MYC*-responsive luciferase reporter was transfected 24 h before the sRNAs. ChIP-chip data and DNase sensitivity profiles from HeLa and HepG2 cells were extracted from the UCSC ENCODE (Encyclopedia of DNA Elements) database. Promoters were grouped depending on the ChIP-chip signal intensity, and the number of PASRs in each group was determined.

Received 16 September 2008; accepted 2 January 2009.

Published online 25 January 2009.

1. Birney, E. *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
2. Shiraki, T. *et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. USA* **100**, 15776–15781 (2003).
3. Kapranov, P., Willingham, A. T. & Gingeras, T. R. Genome-wide transcription and the implications for genomic organization. *Nature Rev. Genet.* **8**, 413–423 (2007).
4. Kapranov, P. *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484–1488 (2007).
5. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141 (2008).
6. Eddy, S. The UCSC Genome Browser ([http://www.genome.ucsc.edu/cgi-bin/hgTables?db=hg18&hgta\\_group=genes&hgta\\_track=rnaGene&hgta\\_table=rnaGene&hgta\\_doSchema=describe+table+schema](http://www.genome.ucsc.edu/cgi-bin/hgTables?db=hg18&hgta_group=genes&hgta_track=rnaGene&hgta_table=rnaGene&hgta_doSchema=describe+table+schema)) (2006).

7. Yang, J. H. et al. snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. *Nucleic Acids Res.* **34**, 5112–5123 (2006).
8. Huttenhofer, A. et al. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* **20**, 2943–2953 (2001).
9. Kawaji, H. et al. CAGE Basic/Analysis Databases: the CAGE resource for comprehensive promoter analysis. *Nucleic Acids Res.* **34**, D632–D636 (2006).
10. Carninci, P. et al. Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genet.* **38**, 626–635 (2006).
11. Berget, S. M. Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**, 2411–2414 (1995).
12. Olofsson, S. O. & Boren, J. Apolipoprotein B: a clinically important apolipoprotein which assembles atherogenic lipoproteins and promotes the development of atherosclerosis. *J. Intern. Med.* **258**, 395–410 (2005).
13. Rada-Iglesias, A. et al. Whole-genome maps of USF1 and USF2 binding and histone H3 acetylation reveal new aspects of promoter structure and candidate genes for common human disorders. *Genome Res.* **18**, 380–392 (2008).
14. Bochnig, P., Reuter, R., Bringmann, P. & Luhrmann, R. A monoclonal antibody against 2,2,7-trimethylguanosine that reacts with intact, class U, small nuclear ribonucleoproteins as well as with 7-methylguanosine-capped RNAs. *Eur. J. Biochem.* **168**, 461–467 (1987).
15. Morris, K. V., Chan, S. W., Jacobsen, S. E. & Looney, D. J. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **305**, 1289–1292 (2004).
16. Ting, A. H., Schuebel, K. E., Herman, J. G. & Baylin, S. B. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. *Nature Genet.* **37**, 906–910 (2005).
17. Janowski, B. A. et al. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nature Chem. Biol.* **3**, 166–173 (2007).
18. Mattick, J. S. RNA regulation: a new genetics? *Nature Rev. Genet.* **5**, 316–323 (2004).
19. Mattick, J. S. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays* **25**, 930–939 (2003).
20. Willingham, A. T. et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309**, 1570–1573 (2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature/](http://www.nature.com/nature/).

**Acknowledgements** We thank L. Cardone, D. Rebolini, M. Kramer, and W. R. McCombie for Illumina sequencing. We wish to thank J. Brosius, J. Schmitz and T. Rozhdestvensky for their help with the small RNA cloning protocol and J. Dumais for technical assistance. K.F.-T. was in part supported by the Schering Foundation. This work was supported in part by grants from the NIH and was performed as part of the ENCODE consortium (G.J.H. and T.R.G.). G.J.H. is an investigator of the Howard Hughes Medical Institute.

**Author Contributions** K.F.-T. and P.K. performed experiments in collaboration with E.D., V.S., R.D. and A.T.W. P.K., S.F., R.S. and G.A. performed data analysis. G.J.H. and T.R.G. planned experiments and wrote the paper.

**Author Information** Sequences generated during this study have been deposited in GEO under accession number GSE14362. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to G.J.H. ([hannon@cshl.edu](mailto:hannon@cshl.edu)) or T.R.G. ([gingeras@cshl.edu](mailto:gingeras@cshl.edu)).

---

#### Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project

**Cold Spring Harbor Laboratory** Katalin Fejes-Toth<sup>1,2\*</sup>, Vihra Sotirova<sup>1,2</sup>, Ravi Sachidanandam<sup>1†</sup>, Gordon Assaf<sup>1,2</sup>, Gregory J. Hannon<sup>1,2</sup>; **Affymetrix** Philipp Kapranov<sup>3\*</sup>, Sylvain Foissac<sup>3</sup>, Aarron T. Willingham<sup>3</sup>, Radha Duttagupta<sup>3</sup>, Erica Dumais<sup>3</sup> & Thomas R. Gingeras<sup>1,3</sup>

<sup>1</sup>Watson School of Biological Sciences, <sup>2</sup>Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA.

<sup>3</sup>Affymetrix, Inc. Santa Clara, California 95051, USA. <sup>†</sup>Present address: Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, 5 East 98th Street, New York, New York 10029, USA.

\*These authors contributed equally to this work.

## LETTERS

# Bidirectional promoters generate pervasive transcription in yeast

Zhenyu Xu<sup>1\*</sup>, Wu Wei<sup>1\*</sup>, Julien Gagneur<sup>1</sup>, Fabiana Perocchi<sup>1</sup>, Sandra Clauðer-Münster<sup>1</sup>, Jurgi Camblong<sup>2</sup>, Elisa Guffanti<sup>3</sup>, Françoise Stutz<sup>3</sup>, Wolfgang Huber<sup>4</sup> & Lars M. Steinmetz<sup>1</sup>

Genome-wide pervasive transcription has been reported in many eukaryotic organisms<sup>1–7</sup>, revealing a highly interleaved transcriptome organization that involves hundreds of previously unknown non-coding RNAs<sup>8</sup>. These recently identified transcripts either exist stably in cells (stable unannotated transcripts, SUTs) or are rapidly degraded by the RNA surveillance pathway (cryptic unstable transcripts, CUTs). One characteristic of pervasive transcription is the extensive overlap of SUTs and CUTs with previously annotated features, which prompts questions regarding how these transcripts are generated, and whether they exert function<sup>9</sup>. Single-gene studies have shown that transcription of SUTs and CUTs can be functional, through mechanisms involving the generated RNAs<sup>10,11</sup> or their generation itself<sup>12–14</sup>. So far, a complete transcriptome architecture including SUTs and CUTs has not been described in any organism. Knowledge about the position and genome-wide arrangement of these transcripts will be instrumental in understanding their function<sup>8,15</sup>. Here we provide a comprehensive analysis of these transcripts in the context of multiple conditions, a mutant of the exosome machinery and different strain backgrounds of *Saccharomyces cerevisiae*. We show that both SUTs and CUTs display distinct patterns of distribution at specific locations. Most of the newly identified transcripts initiate from nucleosome-free regions (NFRs) associated with the promoters of other transcripts (mostly protein-coding genes), or from NFRs at the 3' ends of protein-coding genes. Likewise, about half of all coding transcripts initiate from NFRs associated with promoters of other transcripts. These data change our view of how a genome is transcribed, indicating that bidirectionality is an inherent feature of promoters. Such an arrangement of divergent and overlapping transcripts may provide a mechanism for local spreading of regulatory signals—that is, coupling the transcriptional regulation of neighbouring genes by means of transcriptional interference or histone modification.

To obtain a comprehensive survey of the structure and expression level of transcripts across the yeast genome, we used tiling arrays<sup>3</sup> to profile wild-type transcriptomes in ethanol (YPE), glucose (YPD, SDC) and galactose (YPGal), which together encompass the main laboratory growth conditions of yeast (Supplementary Tables 1 and 2). Transcript start and end positions were mapped to the genome by a segmentation algorithm<sup>16</sup> and subsequent manual curation. To identify CUTs, profiles were measured for a deletion mutant of *RRP6*, coding for an important component of the nuclear exosome, which is involved in the degradation of CUTs<sup>17,18</sup>. Transcripts specific to the *rrp6Δ* mutant were designated as CUTs (Methods). We assigned systematic names to all SUTs and CUTs. Expression profiles are provided in a searchable web database (<http://steinmetzlab.embl.de/NFRsharing>).

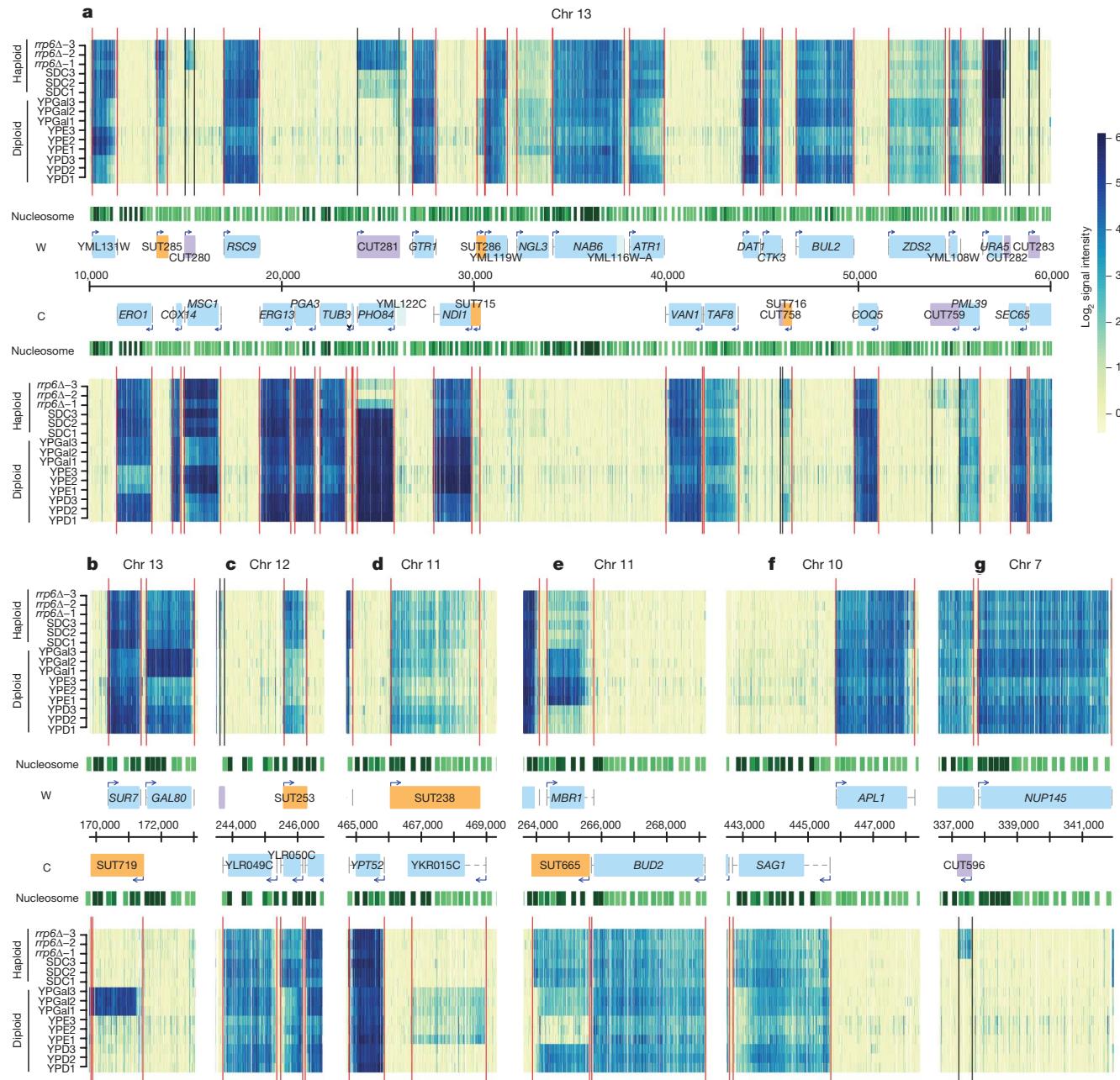
Altogether, 7,272 transcripts were identified, comprising 5,171 verified or uncharacterized open reading frame transcripts (ORFs), 847 SUTs and 925 CUTs (Fig. 1 and Supplementary Table 3). We took advantage of data from different conditions to disambiguate cases of overlapping or immediately adjacent transcripts (Methods). We only used transcripts with confidently mapped 5' ends for analyses involving start sites (5,084 ORFs, 823 SUTs and 704 CUTs; Methods and Supplementary Table 4). For validation, we compared our data to transcript start sites (TSSs) mapped by 5' RACE (rapid amplification of complementary DNA ends)<sup>19</sup>. Eighty-one per cent (1,039 out of 1,281) of TSSs agreed within 50 bases with the 5' RACE results (Supplementary Fig. 1), 3% higher than a recent Solexa sequencing approach<sup>19</sup>. Furthermore, a comparison of our 3' ends with the Solexa data set showed agreement of 61% (2,774 out of 4,551) within 50 bases. In addition, we tested several CUT boundaries and they agreed well with our real-time polymerase chain reaction and 5' RACE validations (Supplementary Fig. 2 and Supplementary Table 5). Altogether, 102 SUTs had a higher expression level in the *rrp6Δ* mutant compared to wild type (Supplementary Table 6), indicating that the distinction between CUTs and SUTs is in some cases condition-dependent, as, for example, the CUT on the antisense strand of *PHO84*, which is stabilized in old cells<sup>11</sup>. CUTs were, overall, shorter (median length 440 bases) than SUTs (median length 761 bases;  $P < 2 \times 10^{-16}$ , Wilcoxon test).

Nucleosome-free promoter regions (or 5' NFRs), which facilitate transcription by allowing RNA polymerase to bind to DNA, have been reported as hallmarks of gene promoters<sup>20–24</sup>. To test whether unannotated transcripts have such hallmarks, we compared our transcript positions with nucleosome maps<sup>22,25</sup>. Consistent with promoter activity at NFRs, all classes of transcripts—ORFs, CUTs and SUTs—showed depletion of nucleosomes upstream of their TSS (Fig. 2a). Furthermore, no nucleosome was detected between 422 out of the 666 (63%) non-overlapping divergent transcript pairs involving at least one unannotated transcript (Methods and Supplementary Table 7). This indicates that these pairs share a single 5' NFR that may function as a bidirectional promoter.

To investigate further the set of potential bidirectional promoters in the yeast genome, we analysed all 1,049 non-overlapping divergent transcript pairs that shared a single 5' NFR. The distribution of distances between their TSSs had an estimated mode (defined as the point with the highest density of the distribution) at 180 bases, whereas their shared NFR lengths had a mode at 131 bases (Fig. 2b). The size of the shared 5' NFRs increased with the inter-transcript distances, in a relationship consistent with a model of a single NFR surrounded by two regions inside the flanking nucleosomes from which transcripts initiate<sup>22,25</sup>.

<sup>1</sup>European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. <sup>2</sup>Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK. <sup>3</sup>Department of Cell Biology, Sciences III, University of Geneva, 30 Quai E. Ansermet, 1211 Geneva 4, Switzerland. <sup>4</sup>European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge CB10 1SD, UK.

\*These authors contributed equally to this work.

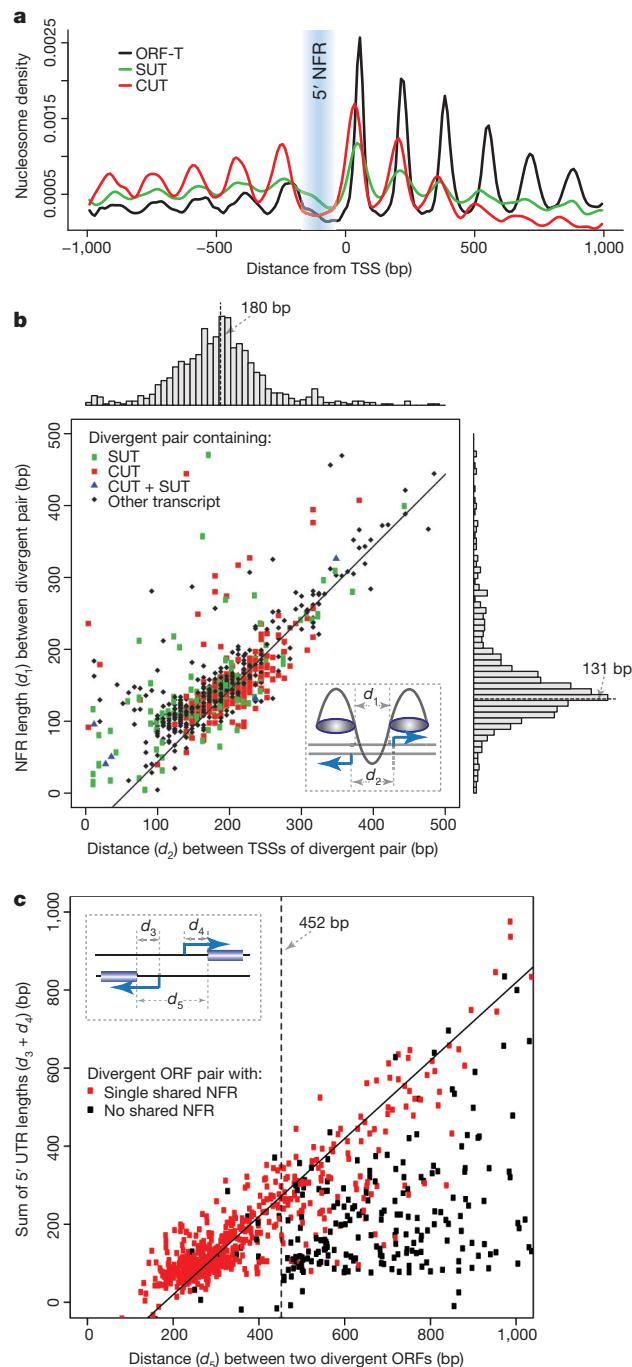


**Figure 1 | Transcript maps.** **a**, Expression data along 50 kb of chromosome 13 (x axis) for the Watson (W, top) and the Crick (C, bottom) strands. (Larger views are available for the whole genome in our searchable web database.) Normalized signal intensities are shown for the profiled samples (y axis): three replicates each for the *rrp6Δ* S96 haploid strain (*rrp6Δ-1-3*), the S96 haploid strain in SDC (SDC1-3), the S1003 diploid strain in YPGal (YPGal1-3) and the S1003 diploid strain in YPE (YPE1-3), and three rows (summarizing nine replicates) for the S1003 diploid strain in YPD (YPD1-3). Vertical lines represent inferred transcript boundaries. Nucleosome positions (green tracks, darker for more significant scores<sup>22</sup>) and genome annotations are shown in the centre: annotated ORFs (blue

boxes) and their mapped UTRs (dashed grey lines), SUTs (orange boxes), CUTs (purple boxes) and transcript start sites (arrows). Coordinates are indicated in base pairs. **b–g**, Examples of transcriptional arrangements; layout is as in **a**. **b**, Tandem gene pair with antisense transcript: *GAL80* shares a NFR with our detected *SUT719*, an antisense transcript of *SUR7*; **c**, antisense transcript *SUT253* originating from both a 5' NFR (of *YLR049C*) and a 3' NFR (of *YLR050C*); **d**, antisense transcript *SUT238* originating from a 5' NFR (of *YPT52*); **e**, *SUT665* originating from a 3' NFR (of *BUD2* and *MBR1*); **f**, divergent promoter of two ORF-Ts with putatively long UTRs; **g**, *CUT596* originating from a 5' NFR (of *NUP145*).

In our analysis, 612 out of 931 non-overlapping divergent protein-coding transcript pairs were found to share a single 5' NFR (66%, Supplementary Table 7). This fraction is considerably higher than the 30% of divergent ORF pairs that were previously estimated to share promoters<sup>26</sup>. Previous studies may have underestimated the number of bidirectional promoters by considering only distances between ORF start codons. Indeed, for divergent ORF-T pairs sharing a 5' NFR (Fig. 2c, red dots), the total untranslated region (UTR) length increased with the distance between the start codons, consistent with a typical size of the inter-transcript distance of a shared promoter being

~180 bases, as evident from Fig. 2b. This relationship holds for a wide range of inter-ORF distances, including cases greater than 1,000 bases, such as *SAG1* and *APL1* (also known as *YAP80*) (Fig. 1f). In contrast, divergent ORF-T pairs separated by multiple NFRs showed no correlation between total UTR length and distance separating start codons (Fig. 2c, black dots). Moreover, most of these pairs were separated by more than 452 bases, which is approximately the minimal size of a region spanned by two NFRs (2 × 131 bases), a nucleosome (146 bases) and two intra-nucleosome regions (2 × 22 bases; Supplementary Fig. 3). These results indicate that bidirectional promoter



**Figure 2 | Properties of divergent transcript pairs. a,** Nucleosome density<sup>22</sup> relative to TSSs, averaged over ORF-Ts (black line), SUTs (green line) and CUTs (red line). **b,** Scatter plot and histograms of shared NFR length ( $d_1$ ) and distance between TSSs ( $d_2$ ) of divergent pairs sharing a 5' NFR. The line corresponds to the regression  $d_1 = d_2 - 2c$ , where the value  $c = 22$  bases was determined from the mode of the distribution of differences between  $d_1$  and  $d_2$ , and corresponds to a typical distance between the NFR and TSS. **c,** Scatter plot of the sum of 5' UTR lengths ( $d_3 + d_4$ ) versus the distance ( $d_5$ ) between coding sequences of divergent ORF-T pairs. The solid line corresponds to the regression  $d_5 = d_3 + d_4 + b$ , where the value  $b = 180$  bases for the typical TSS distance between divergent pairs is taken from **b** above. The vertical dashed line at  $d_5 = 452$  bases is an estimate of the minimal distance for two ORFs to have separate NFRs.

usage is frequent for divergent transcript pairs involving unannotated transcripts and protein-coding genes in any combination.

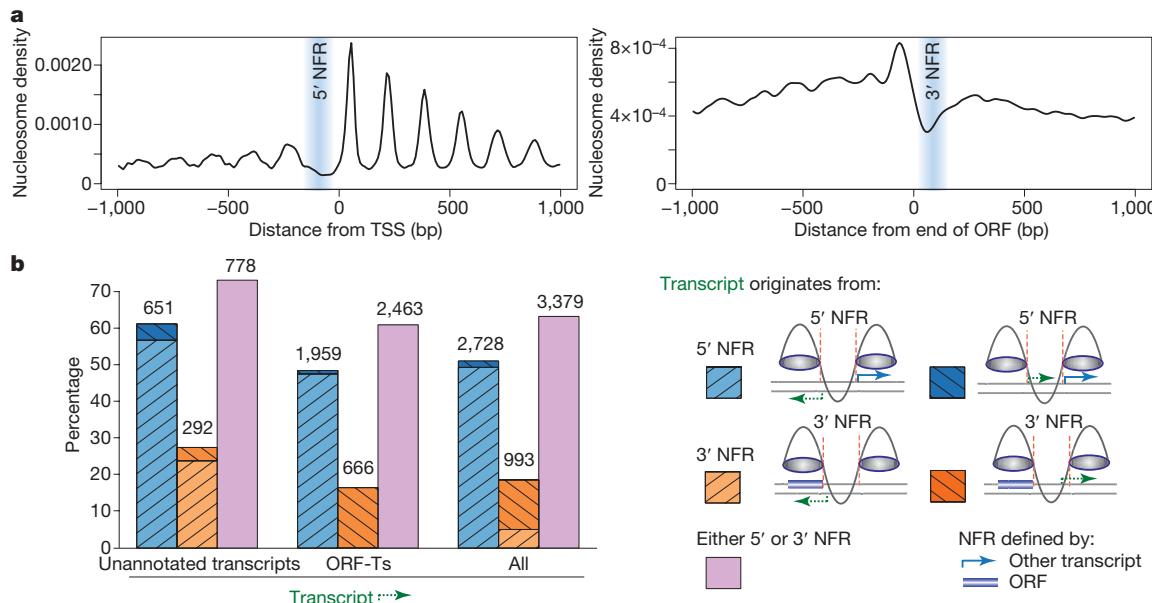
To determine how many of the 5' NFRs initiate transcripts bidirectionally, we selected all nucleosome-depleted regions longer than 80 bases immediately upstream of TSSs, defining a set of 3,965 5' NFRs

(Methods and Supplementary Fig. 4). Of these, 1,318 (33%) were bidirectional, involving half of all transcripts with a mapped 5' NFR (2,656 of 5,339, Supplementary Tables 8–10). The sequences of NFRs detected as bidirectional promoters did not differ significantly from the other 5' NFRs in content of palindromic sequences or GC nucleotides. Among transcripts with mapped 5' NFRs, 61% of unannotated transcripts and 48% of protein-coding transcripts initiated bidirectionally from shared 5' NFRs rather than initiating from their own promoters (Fig. 3b). Of the unannotated transcripts, 90% shared the 5' NFR with a protein-coding transcript. These results indicate that bidirectionality is an inherent property of promoters. In addition to bidirectional transcription, a small number of transcripts was found to initiate in tandem orientation from shared 5' NFRs (Fig. 3b). This number is probably underestimated, however, because of the difficulty of distinguishing immediately adjacent tandem transcripts by microarray hybridization. Altogether, our results indicate that multiple transcripts often initiate from NFRs at promoters in yeast. Additional transcripts will probably be detected by profiling alternative conditions or mutants other than *rrp6Δ*.

In addition to NFRs at promoters, NFRs downstream of stop codons have been reported for most ORFs and are suspected to have a role in transcription termination as well as in the generation of transcripts antisense to the ORF<sup>22</sup> (Fig. 3a). To characterize better such NFRs, we selected all nucleosome-depleted regions longer than 80 bases immediately downstream of stop codons of all verified and uncharacterized ORFs that we detected expressed, defining a set of 2,616 3' NFRs (Supplementary Table 9). Of these, 827 (32%) initiated a transcript. We observed that 27% of unannotated transcripts with a mapped 5' NFR initiated from the 3' NFR of an ORF (Fig. 3b). Together, 3' and 5' shared NFRs thus accounted for most (73%) SUT or CUT initiation, and for most (61%) ORF-T initiation (Fig. 3b, Supplementary Tables 10 and 11 for a list of all pairs). Altogether, these results show a surprisingly high level of NFR sharing, not only in bidirectional promoters but also in 3' NFRs.

The high level of NFR sharing may explain a large extent of antisense transcription<sup>3</sup>, that is, transcription on opposite strands. Seventy per cent of all antisense transcripts with mapped 5' NFRs initiated from a shared NFR. For example, 269 unannotated transcripts initiating from the 3' NFR of an ORF were transcribed antisense to the ORF (for example, *YLR050C* and *MBR1*, Fig. 1c, e). Another recurrent configuration is an antisense transcript starting from the 5' NFR of a downstream tandem transcript. These configurations associate three transcripts; an example is *GAL80*, the 5' NFR of which initiates a transcript antisense to its upstream gene *SUR7* (Fig. 1b). Notably, the level of *SUR7* was lowest in YPGal medium, in which the *SUR7* antisense transcript and *GAL80* had the highest expression (18 further examples are given in Supplementary Table 12). To generalize these observations, we analysed expression correlations across growth conditions among transcript pairs involving at least one SUT. We observed significant expression anti-correlation between sense–antisense pairs, whereas bidirectional pairs of transcripts showed a tendency for co-expression (Supplementary Fig. 5 and Supplementary Table 13; all  $P$  values  $< 10^{-7}$ , Pearson's product moment correlation test). The anti-correlation between sense and antisense transcripts fits the pattern displayed by individual cases of transcriptional interference or inhibitory histone modifications<sup>10–14,27</sup>.

The extent to which the genome-wide set of unannotated transcripts have a biological role, or are merely transcriptional side products (noise) originating from nucleosome-depleted regions<sup>9</sup>, is unknown. The action of transcription itself can be functional even if the transcription product is not. This is the case, for example, with the transcription of the non-coding RNA *SRG1* and the *IME4* antisense transcript, which mediate transcriptional silencing<sup>13,14</sup>. To explore the conservation of transcription initiation from 5' and 3' NFRs, we profiled the transcriptome of *YJM789* (ref. 28), a highly diverged relative of the laboratory strain *S288c*. In rich media (YPD),



**Figure 3 | 5' and 3' NFR sharing.** **a**, Nucleosome density relative to TSSs, averaged over all transcripts (left panel), and relative to translation stop sites, averaged over all ORF-Ts (right panel). **b**, Transcripts initiating from 5' or 3' NFRs of other transcripts. The first block of bars corresponds to unannotated transcripts (1,063), the second to ORF-Ts (4,039) and the third to all transcripts (5,339) with mapped 5' NFRs. Within each block, the bars

about 50% (380 out of 769) of the SUTs expressed in S288c were also found expressed in YJM789 (Methods). The frequencies with which these 380 conserved SUTs were observed sharing NFRs with other transcripts were similar to those in the overall data set. These results indicate that the interlaced architecture of transcript initiation from 5' and 3' NFRs is conserved between these strains of *Saccharomyces cerevisiae*. Why some of the unannotated transcripts are stable and others are unstable remains to be explored. The parasite *Giardia lamblia* produces an abundance of antisense transcripts originating bidirectionally from promoters<sup>29</sup>; consistent with our *rrp6Δ* results, its genome lacks orthologues to several nuclear exosome components. Likewise, the function of bidirectional transcription requires further exploration. One hypothesis is that bidirectional transcription has a role in maintaining an open chromatin structure at promoters. In other instances the combined action of bidirectional promoters and transcriptional regulation by these transcripts, or their generation, may provide a mechanism to spread transcriptional regulatory signals locally in the genome.

## METHODS SUMMARY

Complementary DNA for hybridization was prepared using random- or random-plus-oligo-dT-primer with the addition of actinomycin D during reverse transcription<sup>30</sup>. The hybridization data were normalized and segmented using the Bioconductor package ‘tilingArray’<sup>16</sup>. Segments were then manually curated. Further details can be found in Methods and Supplementary Information.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 13 September; accepted 19 December 2008.

Published online 25 January 2009.

- Bertone, P. et al. Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242–2246 (2004).
- Carninci, P. et al. The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005).
- David, L. et al. A high-resolution map of transcription in the yeast genome. *Proc. Natl Acad. Sci. USA* **103**, 5320–5325 (2006).
- Dutrow, N. et al. Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA-DNA hybrid mapping. *Nature Genet.* **40**, 977–986 (2008).
- Li, L. et al. Genome-wide transcription analyses in rice using tiling microarrays. *Nature Genet.* **38**, 124–129 (2006).
- Stolc, V. et al. A gene expression map for the euchromatic genome of *Drosophila melanogaster*. *Science* **306**, 655–660 (2004).
- Wilhelm, B. T. et al. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239–1243 (2008).
- Kapranov, P., Willingham, A. T. & Gingeras, T. R. Genome-wide transcription and the implications for genomic organization. *Nature Rev. Genet.* **8**, 413–423 (2007).
- Struhl, K. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Struct. Mol. Biol.* **14**, 103–105 (2007).
- Berretta, J., Pinskaya, M. & Morillon, A. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in *S. cerevisiae*. *Genes Dev.* **22**, 615–626 (2008).
- Camblong, J. et al. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**, 706–717 (2007).
- Bird, A. J., Gordon, M., Eide, D. J. & Winge, D. R. Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts. *EMBO J.* **25**, 5726–5734 (2006).
- Hongay, C. F., Grisafi, P. L., Galitski, T. & Fink, G. R. Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. *Cell* **127**, 735–745 (2006).
- Martens, J. A., Laprade, L. & Winston, F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* **429**, 571–574 (2004).
- Birney, E. et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
- Huber, W., Toedling, J. & Steinmetz, L. M. Transcript mapping with high-density oligonucleotide tiling arrays. *Bioinformatics* **22**, 1963–1970 (2006).
- Davis, C. A. & Ares, M. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **103**, 3262–3267 (2006).
- Wyers, F. et al. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737 (2005).
- Nagalakshmi, U. et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).
- Lee, W. et al. A high-resolution atlas of nucleosome occupancy in yeast. *Nature Genet.* **39**, 1235–1244 (2007).
- Shivaswamy, S. et al. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* **6**, e65 (2008).
- Mavrich, T. N. et al. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**, 1073–1083 (2008).
- Whitehouse, I., Rando, O. J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).
- Yuan, G. C. et al. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**, 626–630 (2005).
- Albert, I. et al. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**, 572–576 (2007).
- Hermsen, R., ten Wolde, P. R. & Teichmann, S. Chance and necessity in chromosomal gene distributions. *Trends Genet.* **24**, 216–219 (2008).

27. Uhler, J. P., Hertel, C. & Svejstrup, J. Q. A role for noncoding transcription in activation of the yeast *PHO5* gene. *Proc. Natl Acad. Sci. USA* **104**, 8011–8016 (2007).
28. Wei, W. *et al.* Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc. Natl Acad. Sci. USA* **104**, 12825–12830 (2007).
29. Teodorovic, S., Walls, C. D. & Elmendorf, H. G. Bidirectional transcription is an inherent feature of *Giardia lamblia* promoters and contributes to an abundance of sterile antisense transcripts throughout the genome. *Nucleic Acids Res.* **35**, 2544–2553 (2007).
30. Perocchi, F., Xu, Z., Clauder-Münster, S. & Steinmetz, L. M. Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Res.* **35**, e128 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature/](http://www.nature.com/nature/).

**Acknowledgements** We thank A. Akhtar, A. Ladurner, S. Blandin, R. Aiyar, E. Mancera and E. Fritsch for comments on the manuscript, J. Toedling for discussion and for the template of the website, C. Girardot for data submission to ArrayExpress, N. Proudfoot for access to experimental equipment, and the

contributors to the Bioconductor ([www.bioconductor.org](http://www.bioconductor.org)) and R (<http://www.r-project.org>) projects for their software. This work was supported by grants to L.M.S. from the National Institutes of Health and Deutsche Forschungsgemeinschaft, by a SystemsX fellowship to E.G., by a Roche fellowship to J.C. and by grants to F.S. from SNF and NCCR Frontiers in Genetics.

**Author Contributions** L.M.S., Z.X. and W.W. designed the research; Z.X. and W.W. annotated the transcripts with the help of J.G. and F.P.; W.W. and Z.X. performed analysis of the transcripts with the help of J.G.; F.P. and S.C.-M. performed the array hybridizations; J.C. E.G. and F.S. provided samples for the *rrp6Δ* mutant, and designed and performed validation polymerase chain reaction with reverse transcription and 5' RACE experiments; L.M.S., J.G., F.S. and W.H. supervised the research; and L.M.S., Z.X., W.W., J.G. and W.H. wrote the manuscript.

**Author Information** Raw data are available from ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) under accession number E-TABM-590. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to L.M.S. ([larsms@embl.de](mailto:larsms@embl.de)).

## METHODS

**Strains and media.** *S. cerevisiae* strains used in this work were isogenic to either S288c or YJM789 (Supplementary Table 1). Strains were grown to mid-exponential phase ( $D_{600} \sim 1.0$ ) in YPD (2% peptone, 1% yeast extract, 2% dextrose), YPGal (2% peptone, 1% yeast extract, 2% galactose), YPE (2% peptone, 1% yeast extract, 2% ethanol) or synthetic complete (SDC) medium (0.67% yeast nitrogen base without amino acids, with ammonium sulphate, 2% dextrose and amino acid supplements) (Supplementary Table 2).

**RNA extraction and hybridization to arrays.** Total RNA was extracted from yeast cultures with standard hot phenol protocol and processed for array hybridizations as described previously<sup>30</sup> (Supplementary Methods). Importantly, to remove reverse transcription artefacts, first-strand cDNA was synthesized in the presence of 6.25  $\mu\text{g ml}^{-1}$  actinomycin D. The labelled cDNA samples were denatured and processed for hybridizations<sup>30</sup>. Our analysis is based on replicate hybridizations (Supplementary Table 2).

**Genome sequence and annotation.** Sequence and feature files (.gff files) for S288c were obtained from the *Saccharomyces* Genome Database on 4th September 2007.

**Array data analysis.** Arrays profiled in conditions YPD, YPE and YPGal were normalized with genomic DNA as in ref. 16. Only the probes matching exactly and uniquely to the S288c genome were considered further. The normalized data were jointly segmented using a segmentation algorithm<sup>16</sup> and the automatically identified segments were curated using a custom web-interface (Supplementary Information). This defined the set of manually curated transcripts.

To identify CUTs, arrays for the *rrp6Δ* strain were segmented jointly with the arrays of the wild-type strain in the same condition (SDC). YJM789 arrays were normalized with YJM789 genomic DNA as a reference. Only the probes matching exactly and uniquely to the S288c-aligned part of the YJM789 sequence were considered further. The normalized data were segmented based on the alignment between S288c and YJM789 (ref. 28).

**Transcript categorization.** The manually curated transcripts were overlapped with the genome annotation features and classified as: (1) SUTs, if they did not overlap with existing annotation; (2) ORF-Ts, if they overlapped with a verified or uncharacterized ORF; or (3) other. Transcripts detected solely in *rrp6Δ* were defined as (4) CUTs (see next section). We refer to the union of SUTs and CUTs also as unannotated transcripts. (5) Antisense transcripts were defined as unannotated transcripts that overlapped with other transcripts on the opposite strand.

**Definition of CUTs.** The automatically detected segments for the *rrp6Δ* strain were overlapped with the manually curated transcripts. We defined three criteria: to not overlap any annotated feature; to show higher than twofold expression in *rrp6Δ* compared to wild type; and to be at least 100 bases long. Two types of CUTs were defined. CUTs of the first type were *rrp6Δ* segments that did not overlap any manually curated segments and fulfilled all three criteria. CUTs of the second type were derived from the *rrp6Δ* segments overlapping manually

curated transcripts in either a one-to-one or a many-to-one relationship. The *rrp6Δ*-specific (non-overlapping) parts of these segments were classified as CUTs if they fulfilled all criteria.

**Classification of transcript ends.** Ends of transcripts can be unambiguously detected from the microarray signal when they are not adjacent to another transcript with a higher signal. We classified all transcript ends as being mapped or unmapped. Adjacent transcript ends on a same strand and separated by a distance shorter than 100 bases were investigated as potentially unmapped ends. In such configurations, the 5' end of the downstream transcript was classified as unmapped if all the following criteria were fulfilled: the signal in the intergenic region between the two adjacent transcripts was above background in all conditions; the expression difference between the intergenic region and the downstream transcript was less than twofold in all conditions; and the expression of the downstream transcript was lower than the expression of the upstream transcript signal by twofold in all conditions. Indeed, if any of these three criteria was violated, we considered this as evidence for a transcript starting from this boundary, and considered the 5' end mapped. An analogous definition was applied for the 3' end of the upstream transcript.

**Categorization of adjacent transcript pairs.** To detect adjacent transcript pairs, transcripts were sorted according to the minimum of their start and end positions. Two consecutive transcripts were considered as adjacent pairs. The adjacent pairs were further classified as divergent if the first transcript was on the Crick strand and the second on the Watson strand, as convergent if the reverse was true, and as tandem if both transcripts were on the same strand. To estimate the mode of a (distance) distribution, we used the midpoint of the shorth (the shortest interval that covers half the values).

**Nucleosome data analysis.** The transcripts were compared to the nucleosome map combining the H2A.Z and H3/H4 data from <http://atlas.bx.psu.edu> (refs 22, 25). Two transcripts were considered as sharing a 5' NFR if there was no nucleosome peak between their TSSs. The 5' NFR was defined as the nucleosome-depleted region (at least 80 bases long, see below) immediately upstream of the TSS, and the 3' NFR as the nucleosome-depleted region (at least 80 bases long) downstream of the stop codon of all verified or uncharacterized ORFs. The cutoff value of 80 bases was chosen on the basis of the nucleosome distance distribution. The nucleosome distance distribution showed two modes: one presumably corresponding to the normal nucleosome linker region (18 bases) and a second mode at around 130 bases corresponding to the NFRs (Supplementary Fig. 4).

**YJM789 comparison.** The SGD annotation was first converted into an alignment coordinate system between S288c and YJM789 (ref. 28). The YJM789 transcripts were categorized in the same manner as the manually verified transcripts from S288c-derived strains. S288c SUTs were also mapped into alignment coordinates and overlapped with the unannotated transcripts from YJM789. A transcript was considered expressed in both S288c-derived and YJM789 genomes if the overlap was at least 50% of the transcript lengths measured in the S288c genome.

## LETTERS

# Widespread bidirectional promoters are the major source of cryptic transcripts in yeast

Helen Neil<sup>1</sup>, Christophe Malabat<sup>1</sup>, Yves d'Aubenton-Carafa<sup>2</sup>, Zhenyu Xu<sup>3</sup>, Lars M. Steinmetz<sup>3</sup> & Alain Jacquier<sup>1</sup>

Pervasive and hidden transcription is widespread in eukaryotes<sup>1–4</sup>, but its global level, the mechanisms from which it originates and its functional significance are unclear. Cryptic unstable transcripts (CUTs) were recently described as a principal class of RNA polymerase II transcripts in *Saccharomyces cerevisiae*<sup>5</sup>. These transcripts are targeted for degradation immediately after synthesis by the action of the Nrd1-exosome-TRAMP complexes<sup>6,7</sup>. Although CUT degradation mechanisms have been analysed in detail, the genome-wide distribution at the nucleotide resolution and the prevalence of CUTs are unknown. Here we report the first high-resolution genomic map of CUTs in yeast, revealing a class of potentially functional CUTs and the intrinsic bidirectional nature of eukaryotic promoters. An RNA fraction highly enriched in CUTs was analysed by a 3' Long-SAGE (serial analysis of gene expression) approach adapted to deep sequencing. The resulting detailed genomic map of CUTs revealed that they derive from extremely widespread and very well defined transcription units and do not result from unspecific transcriptional noise. Moreover, the transcription of CUTs predominantly arises within nucleosome-free regions, most of which correspond to promoter regions of bona fide genes. Some of the CUTs start upstream from messenger RNAs and overlap their 5' end. Our study of glycolysis genes, as well as recent results from the literature<sup>8–11</sup>, indicate that such concurrent transcription is potentially associated with regulatory mechanisms. Our data reveal numerous new CUTs with such a potential regulatory role. However, most of the identified CUTs corresponded to transcripts divergent from the promoter regions of genes, indicating that they represent by-products of divergent transcription occurring at many and possibly most promoters. Eukaryotic promoter regions are thus intrinsically bidirectional, a fundamental property that escaped previous analyses because in most cases divergent transcription generates short-lived unstable transcripts present at very low steady-state levels.

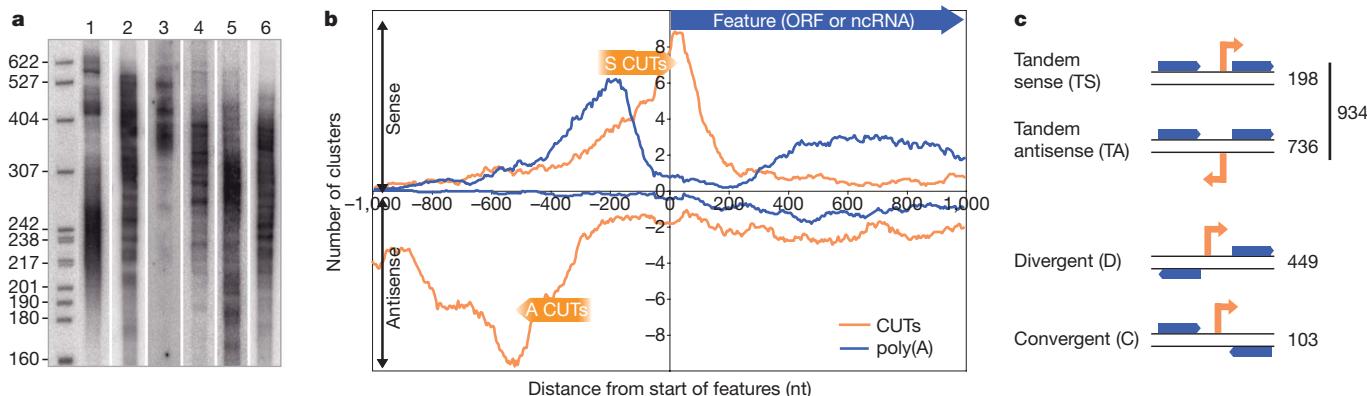
To gain insight on the possible role of CUTs, we determined their complete genomic organization. Because CUTs are capped and mainly degraded by the nuclear exosome and the TRAMP (Trf4-Air1/Air2-Mtr4 polyadenylation) complexes<sup>5</sup>, an RNA fraction highly enriched for CUTs (CUT fraction) was prepared by tandem affinity purification (TAP<sup>12</sup>; see Methods) from a strain in which (1) Cbp20 (also known as Cbc2), a component of the nuclear cap-binding complex, was TAP-tagged and (2) two components of the nuclear degradation machineries were missing (Rrp6 and Trf4 (also known as Pap2)). Supplementary Fig. 1 shows that the model CUT NEL025c was enriched several thousand fold in this fraction relative to wild-type total RNA. To analyse these transcripts, we used an improved 3' Long-SAGE technique<sup>13</sup> (adapted to 454 Roche pyro-sequencing) because it allowed the characterization of the highly heterogeneous 3' ends of the CUTs that result from their Nrd1-dependent transcription

termination<sup>6,7</sup> as well as the distinction between overlapping yet distinct transcripts, which proved critical for the detection of sense CUTs that often overlap mRNA 5' ends (see below). Two barcoded libraries were constructed, the first one from a wild-type poly(A) RNA fraction as a control and the second one from the CUT fraction (polyadenylated *in vitro* for 3' SAGE synthesis), giving rise to 48,118 and 67,022 tags, respectively (see Supplementary Fig. 2 and Methods).

The distribution of the sequenced 3' SAGE tags along the yeast genome was non-random ( $P$  value  $< 10^{-16}$  for both libraries, chi-squared test), but instead highly organized in clusters (see Supplementary Methods for cluster determination and Supplementary Table 3 for a detailed description of the clusters). Analysis of these clusters (Supplementary Fig. 3) indicates that CUTs are distinct from mRNAs yet derive from well-defined transcription units and not from random transcriptional noise.

We defined 1,779 clusters with at least four tags from the CUT fraction (class I, as defined in Supplementary Fig. 3). Most clusters (1,496; 'CUTs' in Supplementary Table 3) do not correspond to annotated features (open reading frames (ORFs) and non-coding RNAs (ncRNAs)). Northern blot analyses of a few revealed heterogeneous transcripts with a small average size of between 200 and 500 nucleotides (nt; Fig. 1a) and Nrd1-dependent transcription termination (Supplementary Fig. 4)—two characteristics of previously known CUTs<sup>5–7</sup>. Among the remaining clusters from this class, 106 could be assigned to non-coding RNA precursors, and 134 clusters, located within intron-containing pre-mRNAs, probably represent degradation products of these transcripts, as suggested from previous studies<sup>14,15</sup> ('ncRNAs' and 'pre-mRNAs' in Supplementary Table 3). Forty-three clusters remained unclassified (mostly found within repetitive elements; 'others' in Supplementary Table 3). Transcription start sites (TSSs) were precisely mapped for 68 of the CUTs (Supplementary Table 4 combining 5' RACE data from Supplementary Fig. 5 and RNase H experiment data from Fig. 2, Supplementary Fig. 6 and data not shown). The 5' ends of these CUTs were usually heterogeneous, with multiple TSSs. Taking the most distal TSS identified as the reference for the 5' end and the maximum 3' SAGE density as the reference for the 3' end positions, it gave a mean size of  $258 \pm 89$  nt ( $\pm$  standard deviation) for the CUTs. In addition, to complete the global visualization of these transcripts, we used the same CUT and poly(A) RNA fractions to perform tiling array hybridizations with a modified procedure adapted from ref. 16 (see Supplementary Methods). These data validated the identification of CUTs performed with the SAGE experiment and confirmed that these transcripts are short. However, technical reasons, specific to the nature of the CUT samples, prevented the use of these data for global statistical analyses (see Supplementary Methods). These data are thus provided for illustration purposes only and can be visualized, together with the SAGE data, at <http://www.pasteur.fr/>

<sup>1</sup>Institut Pasteur, Unité de Génétique des Interactions Macromoléculaires, CNRS, URA2171, 75015 Paris, France. <sup>2</sup>Centre de Génétique Moléculaire, CNRS, Allée de la Terrasse, 91198 Gif-sur-Yvette, France. <sup>3</sup>European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany.



**Figure 1 | Genome-wide analyses of 3' SAGE-tag clusters.** **a**, Northern blot analyses of new CUTs. RNAs from Cbp20-TAP purification (100 ng), performed with the *rrp6Δ dep1-trf4* strain (LMA587) grown in YPD supplemented with doxycyclin, were separated on a 5% polyacrylamide gel and analysed with different riboprobes: (1) CGR088wTa2, (2) CER176wTa3-D, (3) CLR263wTa3-B, (4) CDR213wTa3-E, (5) CDR164cD2 and (6) CGL235wTa2 (primers are listed in Supplementary Table 2). Marker sizes are in nucleotides. **b**, Distribution of tag clusters relative to the start of features. The distances between the start of features (start codons for ORFs) and the nearest CUT 3' SAGE-tag clusters (orange curve) or mRNA 3' SAGE-tag clusters (blue curve) were computed. The figure shows the smoothed distribution of these distances, on either DNA strand, centred on

the start of feature. The orange arrows labelled S CUTs and A CUTs symbolize sense and antisense CUTs, respectively, relative to the downstream feature (blue arrow). Note that the peak of the blue curve in the sense direction corresponds to the ends of tandem upstream mRNAs. **c**, Classification of CUT clusters according to the type of associated genomic intergenic regions. The genomic organization of CUTs (orange arrows) relative to surrounding features (in blue) is schematized. The number of CUTs in each of the four categories defined is indicated on the right. Note that 10 out of the 1,496 clusters were far from an intergenic region and appear to initiate within genes, in antisense (called ‘internal antisense’ in Supplementary Table 3), and are not present on the scheme.

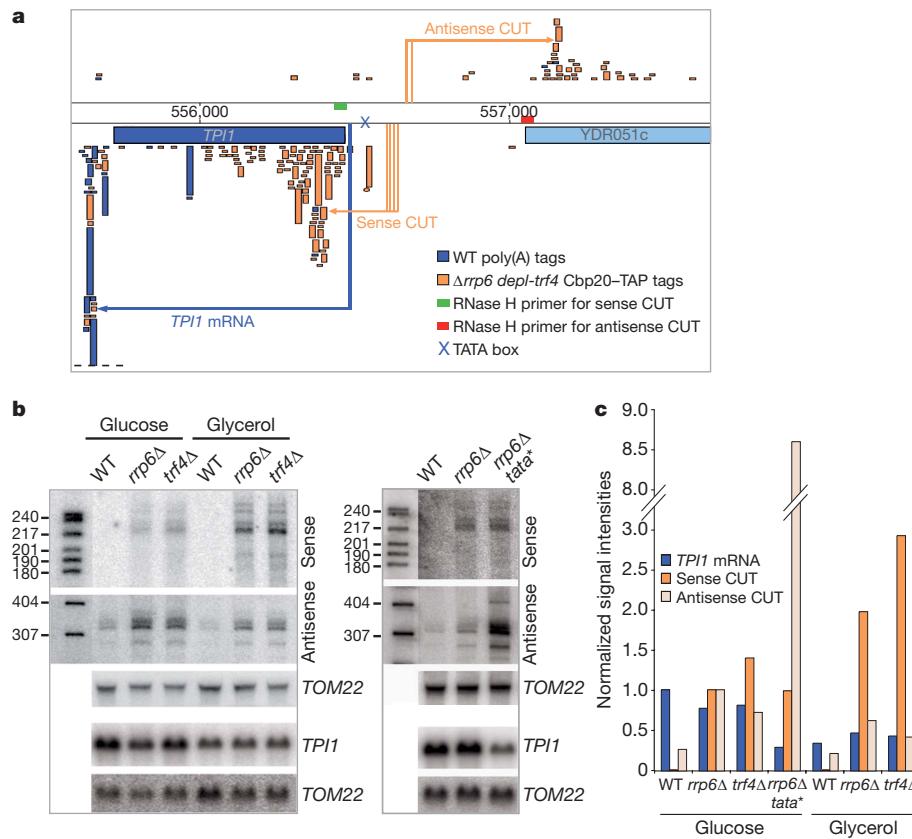
recherche/unites/Gim/genepage/sage.html. Notably, the number of CUTs newly identified was comparable to the number of mRNAs identified from the poly(A) fraction using the same criteria and with an equivalent number of sequences (Supplementary Figs 3 and 7), indicating that CUT transcription units are widespread in the yeast genome, possibly to an extent similar to mRNAs, although their transcription level is probably globally lower.

To study the overall genomic organization of the CUTs, we analysed the distribution of the CUT 3' SAGE clusters relative to the start and end of annotated genomic features (ORFs and ncRNAs). As expected, the positions of mRNA 3' SAGE clusters correlated with the end of features (stop codons for the mRNAs; Supplementary Fig. 8a). The distribution of the CUT 3' SAGE clusters was, in contrast, highly correlated with the start of features, showing a strong peak centred close to the start in the sense orientation and another peak about 550 nt upstream from the start in the antisense orientation (Fig. 1b). Given the average size of ~250–300 nt for the CUTs, this indicates that CUTs preferentially initiate, in either direction, a few hundred nucleotides upstream from genes and thus from intergenic regions (orange arrows in Fig. 1b). Accordingly, as schematized in Fig. 1c, convergent intergenic regions generated CUTs less frequently (6.9%) than divergent intergenic regions (30.2%) or tandemly arranged intergenic regions (62.9%). Most importantly, CUTs transcribed from between tandem features were almost four times more abundant in the antisense (TA: 736 CUTs) than in the sense orientation (TS: 198 CUTs) relative to the flanking genes. This strongly indicates that, in most cases, the promoter of the downstream gene gives rise to a divergent (antisense) transcript in addition to its normal mRNA (see below).

Figure 1b shows that the termination of a large number of sense CUTs overlaps the promoter and the 5' untranslated region (UTR) of genes, having the potential to interfere with and regulate their expression. Several previously unknown regulation mechanisms involving small RNAs that have all the characteristics of CUTs have recently been described. One example is the autoregulation of *NRD1* (ref. 17). In this system, the overexpression of Nrd1 induces the premature termination of the mRNA at an upstream Nrd1-dependent termination site cluster. The resulting unstable small RNA is efficiently detected in our study (Supplementary Fig. 9a). Importantly, in this type of regulation, the CUT and the mRNA are believed to share

the same TSS. It is likely that other CUTs identified here reflect the same type of mechanism suppressing the expression of other genes. However, in most cases that we analysed by RACE or RNase H experiments, TSSs upstream from the mRNA TSS could be specifically detected in the CUT fraction. We thus predict that CUTs sharing their TSS with the mRNA, as exemplified by *Nrd1*, do not constitute the most frequent class. We and others<sup>8,11</sup> have recently shown that several genes of the nucleotide biosynthesis pathway (*URA2*, *URA8*, *IMD2*, *IMD3*, *ADE12*) are regulated by a previously unknown mechanism in which transcription of the mRNAs and the upstream CUTs depends on the same TATA box. Under repressive conditions, transcription initiation mainly starts from proximal sites, generating the CUTs and repressing transcription initiation from the distal TSSs, producing the mRNAs (corresponding CUTs can be visualized in Supplementary Fig. 9b). Yet another regulatory mechanism involving small unstable transcripts is transcription interference, as first described for the *SRG1/SER3* locus<sup>10</sup>. Once again, the non-coding *SRG1* RNA associated with this regulation is readily detected as a bona fide CUT in our study (Supplementary Fig. 9c).

Our data reveal that promoters associated with strongly expressed TS CUTs are enriched for genes involved in glucose catabolism (*P* value of  $2 \times 10^{-9}$  when taking CUTs with at least 20 tags). Three genes from this pathway, *TPII*, *GPM1* and *FBA1*, were chosen for further analyses. The TSSs of their associated TS CUTs were mapped between about one- and two-hundred nucleotides upstream from their mRNA TSSs. Notably, these CUTs were antiregulated relative to their associated mRNAs, being induced in glycerol and repressed in glucose (Fig. 2 and Supplementary Fig. 6). Mutation of the *TPII* TATA box<sup>18</sup> affected expression of the mRNA but not of the sense CUT, as expected because this element is located downstream of the CUT TSS (Fig. 2). This indicates that two independent pre-initiation complexes (PICs) drive expression of the two transcripts, as in the *SRG1/SER3* locus and unlike in the nucleotide biosynthesis pathway. Whether the anticorrelation observed between the expression of the CUTs and the glycolytic mRNAs reflects a causal effect of CUT transcription on mRNA repression remains to be determined. Other types of RNA-mediated regulation have also been described, such as at the *PHO84* locus where an unstable antisense transcript targets repressive chromatin deacetylation to the promoter region of the gene<sup>19</sup>. Our results reveal a number of new loci where transcription



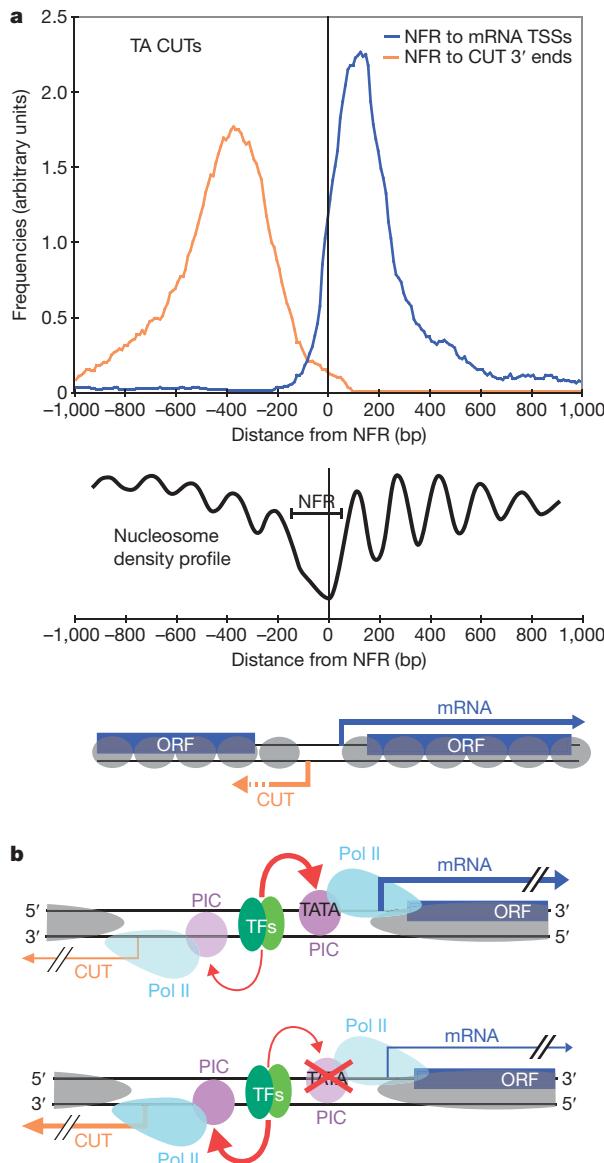
**Figure 2 | Regulation of CUTs associated with *TPI1* mRNAs.** **a**, Summary of transcript organization at the *TPI1* locus. Blue and orange squares correspond to tags obtained by 3' Long-SAGE from wild type (WT) poly(A) RNA (BY4742 strain) and *rrp6Δ depl-trf4* Cbp20-TAP-purified RNA (LMA587 strain), respectively. Units represent position on the chromosome in bp. The blue box represents the *TPI1* ORF, and its corresponding mRNA is indicated by a blue arrow from the TSS, characterized in ref. 25, to the end of the 3' UTR (blue squares). The sense CUT, CDR050cTs2, and the antisense CUT, CDR051cTa3-A, associated with *TPI1* mRNA are represented by an orange arrow, from the TSSs identified below (Supplementary Table 4) and the termination sites (orange squares). **b**, Left: northern blot analysis of *TPI1* mRNA and its associated CUTs in WT (BY4741), *rrp6Δ* (Y01777) and *trf4Δ* (Y06265) strains grown in glucose (YPD) or in glycerol (YPGly). For CUT analysis, total RNA (10 µg) was

treated with RNase H and primers hn213 and hn234, loaded on a 6% polyacrylamide gel and hybridized with the riboprobes hn212–hn214 (sense CUT) and hn235–hn236 (antisense CUT) (Supplementary Tables 2 and 4). Numbers on the left correspond to the molecular ladder weight (units in nucleotides). *TPI1* mRNA was analysed with a DNA probe hn175–hn233 after migration on agarose gel. For both membranes, *TOM22* mRNA was used as a reference (DNA probe cs160–hn310). Right: effects of the *TPI1* TATA-box mutation (*tata\**). As in the left panel but in WT (BY4741), *rrp6Δ* (Y01777) and *rrp6Δ tata\** (LMA808) strains that were grown in YPD. The *TPI1* TATA box TATAAG<sup>18</sup> was changed into TCGCC. **c**, Quantification of the northern blots in **b**. Transcripts were normalized relative to the *TOM22* mRNA. Results correspond to the averaged signals from normalized quantifications from two independent experiments (**b** and data not shown).

of non-coding unstable RNAs could potentially be involved in regulatory mechanisms, and additional work will be required to elucidate the different regulations potentially involved.

In contrast to the sense CUTs, antisense (TA) CUTs do not preferentially associate with particular pathways but are rather ubiquitous, representing by far the most abundant class of CUTs. One simple yet intriguing hypothesis to explain this observation is that TA CUTs result from divergent transcription from gene promoter regions. Such a general bidirectionality of promoters has previously been observed in the simplified eukaryote *Giardia lamblia*<sup>20</sup>, in which an efficient quality control mechanism that degrades CUTs is probably absent. Bidirectionality of promoters would thus reflect an intrinsic property of the eukaryotic transcription machinery but was previously overlooked because the instability of the resulting divergent transcripts prevented, until now, their identification. Fully consistent with this hypothesis, the distribution of TA clusters (but also TS and D clusters, Supplementary Fig. 10a, b) was found to be strongly correlated with the position of the nucleosome-free regions (NFRs) that are most often found associated with gene promoters<sup>21</sup> (Fig. 3a). Such a connection between pervasive transcription and NFRs had been suggested previously<sup>22,23</sup>. The C clusters appeared less associated with NFRs (Supplementary Fig. 10c), possibly as a result of the activity of the Isw2

chromatin remodelling complex that has been shown to suppress the expression of antisense CUTs generated from 3'-end NFRs<sup>24</sup>. Because the TA cluster distribution peaks about 400 nt away from the minimum of nucleosome densities, this indicates that transcription of antisense CUTs predominantly originates within these NFR regions (see scheme in Fig. 3a, bottom). If TA CUTs mostly originate from cryptic divergent transcription from standard gene promoters, it could be expected that the transcription of the TA CUTs and their associated genes would be coordinated. Indeed, at the *TPI1* gene locus, an antisense CUT is also expressed in addition to the TS CUT. Quite notably, this TA CUT is co-regulated with the mRNA, in contrast to the anti-regulated TS CUT (Fig. 2), fully consistent with the hypothesis that transcription of the divergent TA CUT and the mRNA depends on regulation by a common promoter region. It is possible that the efficiency of PIC assembly on either side of the transcription factor binding sites impacts transcriptional directionality. Consistently, although mutation of the *TPI1* TATA did not affect TS CUT transcription (see above), it strongly enhanced transcription of the TA CUT (Fig. 2). This supports the notion that the TA CUT and the mRNA transcripts originate from the same promoter region but from distinct PICs competing for the same pool of general transcription factors and polymerase recruited by the transcription activators (Fig. 3b).



**Figure 3 | Distribution of TA CUTs relative to NFRs.** **a**, Top: out of the 736 TS CUTs, 629 (85%) were found to terminate within  $\pm 1,000$  nt of a NFR. Their distribution relative to the minimum of nucleosome densities of the nearest NFR, zero on the  $x$  axis, is shown in orange. The blue line represents the distribution of distances between the mRNA TSSs and their closest NFRs. Middle: the graph in black shows, on the same scale, a schematic representation of a typical nucleosome density profile upstream from start codons<sup>21</sup>. Bottom: schematic representation, on the same scale, of a typical organization of TA CUTs and associated ORFs deduced from the above distribution profiles. The blue and orange arrows represent the mRNA and CUT transcripts, respectively. The dashed line at the end of the CUT symbolizes 3'-end heterogeneities. ORFs are shown as blue boxes, and nucleosomes as grey ellipses. **b**, Schematized model for TA CUT formation. Symbols are as in **a**. In addition, transcription factors (TFs), RNA polymerase II (pol II) and pre-initiation complexes (PICs) are in green, purple and light blue, respectively. TATA indicates the position of a TATA box. In this model, PIC formation is driven, on either side, by transcription factors. Sequence contexts at the sites of PIC assembly influence the balance between both sides. In normal conditions, the most favourable site for PIC formation (presence of a TATA box, for example) is on the gene side, shifting the balance of PIC formation towards the gene (top) at the expense of other weaker sequences on the opposite side. When the TATA box is mutated (bottom), as in Fig. 2, the balance can be reversed and the pool of general transcription factors recruited by the activators becomes available to form a PIC and initiate pol II transcription on weaker sites in the opposite direction.

Our analyses show that CUTs are extremely widespread in yeast, that a number of CUTs overlap promoters and mRNA 5' UTRs and can potentially interfere with gene expression, and that the most frequent class of CUTs results from cryptic divergent transcription from promoters, which probably give evidence for a fundamental property of the eukaryotic transcription machinery—that is, that promoter regions are intrinsically bidirectional.

## METHODS SUMMARY

The poly(A)<sup>+</sup>-enriched fraction was obtained by purification of total RNA from the BY4742 strain on Dynabeads Oligo (dT)<sub>25</sub> (Dynal, Invitrogen). The CUT fraction was obtained from LMA587 Cbp20-TAP *rrp6Δ tetO7-TRF4* (tet O7-TRF4 is repressed on addition of doxycycline) strain grown in YPD at 25 °C, supplemented with doxycycline for 16 h to deplete Trf4 protein. The TAP method<sup>12</sup> was used to purify RNA associated with Cbp20-TAP protein. To preserve RNA integrity, cells were ground in liquid nitrogen and RNase inhibitors were added (rRNasin, Promega, and ribonucleoside vanadyl complex, New England Biolabs). After TAP, this fraction was poly-adenylated *in vitro* with the poly(A) tailing kit (Ambion).

These RNA fractions were used to construct two 3' Long-SAGE libraries. Different steps of the initially described protocol<sup>13</sup> were improved to enhance efficiency, as detailed in Methods, and the concatenation of the tags was adapted to pyro-sequencing with final products of 150 base pairs long containing 4 tags framed by defined linkers (Supplementary Fig. 2). The tetratags had a specific barcode sequence at their 5' end that distinguished both libraries and allowed their combined 454 Roche sequencing (Genoscope). SAGE-tag identification was performed as described in Methods. The genomic location of tags was determined by a blast search. Clusters of tags were defined by computing tags separated by less than 100 nucleotides.

The same RNA fractions were also hybridized on tiling arrays according to ref. 16, except for the CUT fraction for which, starting with 100 ng of RNA, the GeneCHIP whole-transcript sense target labelling assay (Affymetrix) was applied, with addition of actinomycin D (ActD) during reverse transcription.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 12 September; accepted 24 December 2008.

Published online 25 January 2009.

- Davis, C. A. & Ares, M. Jr. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **103**, 3262–3267 (2006).
- Johnson, J. M., Edwards, S., Shoemaker, D. & Schadt, E. E. Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet.* **21**, 93–102 (2005).
- Kapranov, P. et al. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296**, 916–919 (2002).
- Kapranov, P. et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484–1488 (2007).
- Wyers, F. et al. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737 (2005).
- Arigo, J. T., Eyler, D. E., Carroll, K. L. & Corden, J. L. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell* **23**, 841–851 (2006).
- Thiebaut, M., Kisleva-Romanova, E., Rougemont, M., Boulay, J. & Libri, D. Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1-Nab3 pathway in genome surveillance. *Mol. Cell* **23**, 853–864 (2006).
- Kuehner, J. N. & Brow, D. A. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol. Cell* **31**, 201–211 (2008).
- Martens, J. A., Laprade, L. & Winston, F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* **429**, 571–574 (2004).
- Martens, J. A., Wu, P. Y. & Winston, F. Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes Dev.* **19**, 2695–2704 (2005).
- Thiebaut, M. et al. Futility cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Mol. Cell* **31**, 671–682 (2008).
- Rigaut, G. et al. A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnol.* **17**, 1030–1032 (1999).
- Wei, C. L. et al. 5' Long serial analysis of gene expression (LongSAGE) and 3' LongSAGE for transcriptome characterization and genome annotation. *Proc. Natl. Acad. Sci. USA* **101**, 11701–11706 (2004).
- van Hoof, A., Lennertz, P. & Parker, R. Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.* **20**, 441–452 (2000).
- Bousquet-Antonielli, C., Presutti, C. & Tollervey, D. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102**, 765–775 (2000).

16. David, L. et al. A high-resolution map of transcription in the yeast genome. *Proc. Natl Acad. Sci. USA* **103**, 5320–5325 (2006).
17. Arigo, J. T., Carroll, K. L., Ames, J. M. & Corden, J. L. Regulation of yeast *NRD1* expression by premature transcription termination. *Mol. Cell* **21**, 641–651 (2006).
18. Scott, E. W. & Baker, H. V. Concerted action of the transcriptional activators REB1, RAP1, and GCR1 in the high-level expression of the glycolytic gene *TPI*. *Mol. Cell. Biol.* **13**, 543–550 (1993).
19. Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G. & Stutz, F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**, 706–717 (2007).
20. Teodorovic, S., Walls, C. D. & Elmendorf, H. G. Bidirectional transcription is an inherent feature of *Giardia lamblia* promoters and contributes to an abundance of sterile antisense transcripts throughout the genome. *Nucleic Acids Res.* **35**, 2544–2553 (2007).
21. Lee, W. et al. A high-resolution atlas of nucleosome occupancy in yeast. *Nature Genet.* **39**, 1235–1244 (2007).
22. Ito, T., Miura, F. & Onda, M. Unexpected complexity of the budding yeast transcriptome. *IUBMB Life* **60**, 775–781 (2008).
23. Struhl, K. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Struct. Mol. Biol.* **14**, 103–105 (2007).
24. Whitehouse, I., Rando, O. J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).
25. Nagalakshmi, U. et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature/](http://www.nature.com/nature/).

**Acknowledgements** We thank D. Libri for the gift of strains, D. Libri, C. Thermes, C. Saveanu and M. Fromont-Racine for discussions and for critical reading of the manuscript, staff of the Genoscope (Evry) for sequencing, and A. Doyen, L. Decourty, S. Clauerd-Muenster and T. Bähr-Ivacevic for technical assistance. This work was supported by the Institut Pasteur, the CNRS, the ANR (CUT program), the European Science Foundation (RNA quality program), and the National Institutes of Health and the Deutsche Forschungsgemeinschaft. H.N. is supported by the ANR (CUT program).

**Author Contributions** H.N. performed the experimental work except the sequencing, which was performed at Genoscope, Evry, and the tiling array experiments, which were performed in the group of L.M.S. and analysed by Z.X.; C.M. performed most of the bioinformatics analyses with the help of Y.d'A.-C. for the NFRs analyses; A.J. and H.N. designed the research and A.J. supervised the work; and A.J., H.N. and C.M. wrote the manuscript.

**Author Information** Raw data are available from ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) under accession number E-MTAB-75 for SAGE data and E-TABM-602 for tiling array data. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to A.J. (alain.jacquier@pasteur.fr).

## METHODS

Methods for CUT RNA fraction purification and 3' Long-SAGE tag synthesis, specific to this work, are described here. For additional detailed methods, see Supplementary Information.

**CUT RNA fraction purification by TAP.** Six litres of LMA587 yeast cells (Supplementary Table 1) were grown in YPD at 25 °C up to a  $D_{600}$  of 0.07, at which point doxycycline was added to a final concentration of 10 µg ml<sup>-1</sup> and the culture incubated for a further 16 h at 25 °C. The culture reached a  $D_{600}$  of about 0.6. Cells were then collected, washed in cold water, re-suspended in 12 ml of cold breaking buffer containing 0.1 M Tris-HCl, pH 8, 0.1 M NaCl, 20 mM ribonucleoside vanadyl complex (New England Biolabs), 2× protease inhibitors cocktail (Complete, Roche) and 300 U ml<sup>-1</sup> of rRNasin (Promega), and then frozen as drops in liquid nitrogen. Cells were broken using a cryogenic impact grinder (Freezer Mill 6770, Spex). The purification procedure was adapted from ref. 12. Frozen powder of ground cells was thawed and centrifuged for 45 min at 22,000g (4 °C). Igepal was added to supernatant to a final concentration of 0.1%. IgG Sepharose beads (700 µl of suspension, GE Healthcare) were equilibrated with binding buffer (20 mM Tris-HCl, pH 7.4, 0.1 M NaCl and 0.1% Igepal), added to the total extract and incubated for 2 h at 4 °C on a rotating wheel (low speed). After centrifugation at 100g (4 °C), beads were washed twice by incubation in 10 ml of binding buffer during 10 min at 4 °C on a rotating wheel, and once in 10 ml of binding buffer supplemented with 1 mM DTT. Beads were re-suspended in 600 µl of DTT-binding buffer, 200 U of AcTEV protease (Invitrogen) was added to the suspension, and the reaction was incubated for 2 h at 16 °C on a rotating wheel. The supernatant was recovered and beads were washed twice with 200 µl of binding buffer. Supernatants were pooled and CaCl<sub>2</sub> was added to a final concentration of 2 mM. Calmodulin Sepharose beads (450 µl of beads suspension, GE Healthcare) were equilibrated with binding buffer containing 2 mM CaCl<sub>2</sub>, added to pooled supernatants and incubated during 50 min at 4 °C on a rotating wheel. Beads were recovered and washed three times with 1 ml of cold CaCl<sub>2</sub>-binding buffer. Elution was performed by incubating the beads in 500 µl of elution buffer (20 mM Tris-HCl, pH 8, 50 mM NaCl and 5 mM EGTA) during 5 min at room temperature (25 °C), twice. Pooled supernatants were treated twice with acidic phenol and precipitated with glycogen, NH<sub>4</sub>-acetate and ethanol.

**In vitro poly-adenylation of RNA from the CUT RNA fraction.** In the Trf4-depleted LMA587 culture, CUTs are no longer polyadenylated. They were polyadenylated *in vitro* before the cDNA synthesis that uses oligo-dT priming. The RNA purified above was cleaned on RNeasy columns (Qiagen) and treated with the poly(A) tailing kit from Ambion. RNA, pre-heated for 10 min at 65 °C and cooled on ice, was incubated with *Escherichia coli* poly(A) polymerase (4 U per µg

of RNA) and 50 µM cordycepin 5'-triphosphate (to limit the length of poly(A) tailing) during 1 h at 37 °C.

**Synthesis of 3' Long-SAGE tags.** The following protocol was adapted from ref. 13 and schematized in Supplementary Fig. 2. Double-stranded cDNA was synthesized with SuperScript double-stranded cDNA synthesis kit (Invitrogen). First-strand synthesis was primed with an anchored oligo(dT)<sub>16</sub> primer containing a GsuI restriction site and a 5'-biotin (hn27; Supplementary Table 2). One microgram of RNA and 10 µg of primer were incubated at 65 °C for 10 min and put on ice. The reaction was as described by the manufacturer, except that 0.5 mM methyl-dCTP replaced dCTP and 30 µCi of [ $\alpha$ -<sup>32</sup>P] dATP was added. After incubation for 1 h at 43 °C, second-strand was synthesized as described during 2 h at 16 °C, except that an excess of dCTP was added at a final concentration of 0.6 mM. After purification on Qiagen columns, cDNAs were recovered using streptavidin beads according to manufacturer's protocol (MyOne, Dynal). Biotinylated cDNAs bound on beads were then digested with GsuI (30 U, Fermentas) for 2 h at 30 °C. Supernatant was treated with phenol/chloroform and precipitated. cDNAs were divided in four and independently ligated to linkers A, B, C1 and C2 (Supplementary Table 2, prepared as described<sup>13</sup>, linkers A1/B1 for the WT poly(A) library and A2/B2 for the CUT RNA fraction library). Ligations were performed with 10 µg of annealed linker and 4,000 U of T4 DNA ligase (New England Biolabs) overnight at 16 °C, and subsequently purified on Qiagen columns. Each ligation was digested with 4 U of MmeI (New England Biolabs) for 1 h 30 min at 37 °C, treated with phenol/chloroform and precipitated. Monotags, released by MmeI digestion, were purified on an 8% non-denaturing polyacrylamide gel after visualization by autoradiography. DNA was eluted by incubation of the gel slice in buffer containing 0.3 M NaCl, 0.2 M Tris-HCl, pH 7.5, 25 mM EDTA and 2% SDS, overnight at 37 °C, followed by a phenol/chloroform extraction and a precipitation. Ditags were then obtained by pooling monotags A and C1 together in one hand, and monotags B and C2 on the other hand, and incubating with 4,000 U of T4 DNA ligase (New England Biolabs) at 16 °C overnight. After precipitation, ligations were purified on an 8% non-denaturing polyacrylamide gel after visualization by autoradiography, to only recover the ditags A-C1 and B-C2. DNA was extracted from gel as indicated above, treated with phenol/chloroform and precipitated. Ditags were amplified by PCR, using primers hn31-hn89 and hn31-hn90, respectively, and products were purified on Qiagen columns. To ligate ditags into tetratags, BsaI digestion was first performed. Digested products were purified on 3.5% agarose gel (band of interest is 75 bp +4 nt long) and DNA was recovered with NucleoSpin extract II kit (Macherey-Nagel). Tetratags were obtained by ligation of ditags A-C1 and B-C2 using T4 DNA ligase (New England Biolabs) and amplified by PCR using primers hn89 and hn90.



# MOVERS

**Walter Rosenthal, scientific director,  
Max Delbrück Center for Molecular Medicine,  
Berlin, Germany**



**1996–2008:** Director, Leibniz Institute for Molecular Pharmacology, Berlin, Germany

**1993–96:** Chairman, Rudolf Buchheim Institute for Pharmacology, Giessen, Germany

**1991–93:** Heisenberg fellow and visiting professor, Department of Cell Biology, Baylor College of Medicine, Houston, Texas

Many scientists know their career leanings from an early age. Walter Rosenthal did not. Then, at the age of 19, while driving an ambulance for the German Red Cross, he discovered his vocation in medicine.

Rosenthal is amazed by how far his interest in medicine and pharmacology has taken him. As the new scientific director of the Max Delbrück Center (MDC) for Molecular Medicine in Berlin, he will follow his translational-research mantra: turn basic science into clinical solutions.

While studying medicine at the Justus Liebig University in Giessen, Rosenthal assumed he would become a general practitioner. "I didn't know research was an option when I first started my education," he says. His chats with pharmacologists introduced him to the demanding, open discussions inherent to research. He went on to develop assays to detect cellular signalling molecules as part of a research thesis.

A postdoc spent primarily at the Free University of Berlin would prove crucial. Working with pharmacologist Günter Schultz, Rosenthal began isolating and characterizing G proteins, a large family of cell-surface signalling proteins. "I was lucky to enter the field of G proteins when it was exploding with opportunities," he says.

Rosenthal returned to his medical roots when Mariel Birnbaumer, a colleague at Baylor University in Houston, Texas, invited him to help clone a receptor implicated in diabetes insipidus. Later he identified gene mutations in patients with a congenital form of the disease. Molecular medicine has been Rosenthal's career pursuit ever since.

Although he had just established a lab at the University of Giessen, Rosenthal couldn't resist an invitation to lead the Institute for Molecular Pharmacology (FMP) in Berlin. "I was attracted to the FMP by the idea of having pharmacologists, structural biologists, chemists and molecular geneticists all under one roof," he says.

But first he had to move the former East German institute to its current location near the MDC in Berlin-Buch. Detlev Ganzen, former chief executive of Berlin's Charité hospital, says Rosenthal had the academic standing, demeanour and stature to overcome the political and infrastructure hurdles.

Rosenthal plans to expand the MDC's study of systems biology and forge translational-research collaborations in cardiovascular disease. "There is no recipe to best connect excellent basic research with clinical medicine," he says. "But we have to improve it somehow." ■

**Virginia Gewin**

# BRICKS & MORTAR

## New centres plan for healthy ageing

Competition is heating up for researchers in the rapidly growing field of ageing research. Some two dozen faculty-level scientists and postdocs are being recruited by new or expanding US institutes.

"Ageing is hot because it covers almost every aspect of biology and underlies most major human diseases," says Andrew Dillin, director of the new Center for Aging Research at the Salk Institute in La Jolla, California.

Funded by a \$5-million grant from the Glenn Foundation for Medical Research in Carpinteria, California, the centre will support up to three new research groups. It will build on the Salk Institute's research strengths in metabolism, stem-cell biology, ageing of dividing cells and ageing of organisms. The centre is recruiting seven or eight postdocs to work on a joint project in two labs. The funding will also accommodate two or three visiting scholars each year. Staying for three to six months, they will build collaborative research projects involving multiple labs, and organize a symposium on a topic of their choice.

Collaborations may extend beyond the Salk campus. Although the details aren't finalized, Salk biologist Fred Gage says there is potential for synergy with the Massachusetts Institute of Technology (MIT) and

Harvard University, two other Glenn funding recipients. MIT's Glenn Laboratory for the Science of Aging — which focuses on ageing regulators, such as proteins known as sirtuins that allow cells to survive damage and delay death — is recruiting four postdocs. The \$5-million Glenn gift received last autumn will fund its work on mouse models of ageing and ageing diseases, including Alzheimer's, cancer, osteoporosis and metabolic diseases.

The Buck Institute for Age Research in Novato, California, is seeking a faculty-level stem-cell researcher now and will soon fill 12 more faculty posts at a \$41-million research facility, half-funded by the California Institute for Regenerative Medicine. Scheduled to open in 2010, it will provide more dedicated space for stem-cell biology, an area with potential to diagnose and treat age-related disease. "Ageing is one of the frontiers of science and we've created an institute that sits at the interface between ageing and age-related disease research," says biologist Gordon Lithgow.

"One of the overarching questions is what constitutes healthy ageing," says Gage. "We want to learn how to optimize the process because, unlike a disease, ageing isn't going away." ■

**Virginia Gewin**

### POSTDOC JOURNAL

## It takes a lab to raise a child

My return to work has been smoother than I expected. Why the ease of re-entry? In my first month back I worked flexible hours, many from home. Working half-days in the lab really made these big changes easy for baby and mum, as we built our new schedule in increments rather than all at once. Also, everyone at work has been extremely baby-friendly.

When I needed to pop into the lab to throw samples in the water bath or prep tomorrow's experiment, there was always a colleague happy to play peek-a-boo. I even took my son to a local conference. Most importantly, I credit the patience and support of my stellar supervisor and lab mates, and my partner, who is taking six months' paternity leave so I can return to full-time work.

Why my hurry back to the bench when I could have taken a full year of leave? As might be expected, even six months of down-time has reduced the potential number of manuscripts that I'll have in print this year. In addition to the usual pressure to publish, we, like so many other labs, have major grant deadlines looming. These have no time or tolerance for babies, no matter how cute.

The honeymoon period is certainly over. The next month will be packed with lab work and data analysis. There will be no extra rest for the wickedly sleep-deprived.

**Julia Boughner** is a postdoc in evolutionary developmental biology at the University of Calgary, Canada.

# Penance

An act of faith.

## T. F. Davenport

"For a small fee, I can say a prayer of some kind."

"No thank you, Father."

"Do you have any metal plates in your skull?"

"No, Father."

The blessing takes place in the hallowed small room at the back of the church. It was once a minister's office. The supplicant sits with his spotted pate bowed, baseball cap resting across his thighs. Quietly grateful for the end of a long day, Father Lin applies the wand to the old man's head. He traces the motions of redemption and bliss: tight circles above the right ear.

"Aaaahhh..."

The old man slumps back in the armchair. Father Lin relaxes. There's always a danger that instead of the gentle, kind god advertised and sold by the church, its vengeful Old Testament twin will show up. That happens quite rarely, always by accident, but not — thank goodness — today.

The old man snuffles back tears, fitting the baseball cap on his head. "Thank you, Father." He pulls Father Lin into a hug; church rules oblige him to reciprocate. Holding the man and patting his back, staring over the cap's stiff bill into an Oakland Athletics logo, Lin contemplates using the wand on himself. The end of the week means no one to talk to, no one to pull him into the unwanted hug which, by the end of his day off, he always longs for again.

But he doesn't use the wand on himself. Somehow that would be cheating. He escorts the old man out, locks the door, turns off the neon sign. He files the man's medical waiver then goes to the broom closet. All that remain are a few minutes mopping the scuffed linoleum, then a Starbucks latte on the way home.

The flow of supplicants started early that day, keeping up until closing with hardly a pause. One after another, holding their numbered tickets, they had their minute in the overstuffed armchair. They told him their sins, troubles and grievances against life. He listened without judgement. Some of them cried when he picked up the wand. The wand is a solenoid on a plastic rod, hooked up to a large capacitor; the sacrament is a dime-sized brain region known

to produce epileptic seizures — and, perhaps as a consequence, ecstatic visions. It takes careful work to get the second without the first.

He's dragging the mop across the wood-grain linoleum when he hears someone knocking at the church's front door. Faintly annoyed, he carries on mopping, but whoever it is just won't go away. He leans the mop against the wall and goes to the tall windows flanking the door. The building was once a church, a real church of the sin-and-pray variety. Unable to maintain



it, the dwindling parish had no choice but to sell: either to Abner Lin with his associate's degree in neurodivinity, or to a neighbourhood Muslim group, eager to open an interfaith dialogue and offer classes in adult literacy. There was no deliberation. Father Lin took possession and hauled the pews out onto the kerb for the trash service. Now the only trace of the building's past is the double front door, wide enough for a coffin and pallbearers, and the skinny windows flanking it.

Through this window he now peers at a stooped old woman clutching the handrail. With a kerchief pinned over her grey hair and a tired, desperate expression, she looks like an escapee from a medieval peasant theme park. Radiating utter hopelessness, she knocks at the door again.

He opens it. He means to tell her the

church is closed until Monday, but the light in her eyes upon seeing him, a plain expression that nobody, ever, has been so kind to her, brings a stammer to his lips instead.

"Thank you!" she says. "Thank you!"

"But, ma'am..."

"Thank you," she says firmly. Her eyes brim with gratitude as she takes his arm, and Lin finds himself leading her inside. His heart swells as he takes her through the nave, past the shadow of the absent altar, into the sanctum. He sits her down in the armchair; her tears of despair turn to gratitude, and now the words pour out of her. Polish, he wonders? Whatever the language, Father Lin knows guilt when he hears it. He waits for her confession to end before he gives her the blessing. Perhaps a different blessing from the one she expects, but this one at least is guaranteed to work. As she winds down she's still crying. She holds his hand to her wet cheek. He doesn't have it in him to make her pay.

She signs the medical waiver without looking at it.

Father Lin cranks up the voltage as the woman rattles her hands in prayer. She pauses, briefly alarmed, as he removes her metal hairpins.

The wand hums as he brings it close to her head. She looks up, bumping the wand with her temple. Magnetic fields reach into her skull and in place of the small cortical patch responsible for religious ecstasy, they find her motor cortex. Her hand curls onto itself. Tendons

pop out on her neck as she emits a series of strangled grunts. She slides to the floor and curls there into a fetal convulsing heap. Foamy spittle collects on her lips.

Father Lin curses and rubs his eyes. Then he takes out his phone and stabs the autodial. The paramedics arrive to find him seated and the old woman prostrate, hugging his ankles. His face is worn out with fatigue and loneliness, and the woman tearfully repeats the only English she knows.

"Thank you, thank you..."

**T. F. Davenport** lives in California, studying for a doctorate in cognitive science. In his spare time, he would be writing more science-fiction stories, but he has no spare time, because he's studying for a doctorate in cognitive science.